The international journal of science / 24 October 2019

# nature



## QUANTUM SUPREMACY

Classical supercomputer outperformed by quantum chip for the first time

The international journal of science / 24 October 2019

nature

# A new look for *Nature*

**The journal has been redesigned for clearer research communication in the digital age.**

From today, *Nature* will look a little different. We are unveiling a redesign that will, we hope, help us fulfil our mission to serve researchers and disseminate scientific knowledge worldwide.

This design has been in development for well over a year, and is a much-needed update that helps us – in our 150th year – to communicate science with fresh clarity and style. We love it, and we hope that you, our readers, do, too.

*Nature* has had a number of design transformations over its history – but they were all based on one assumption: that our content would be accessed through the medium of static ink printed on a physical page. Not any more. That's why we have developed a design that is suited to digital platforms – where the vast majority of readers now find us – while at the same time producing a clear and engaging printed edition.

In surveys and interviews, readers told us that our text can be hard to read; and that research articles increasingly need to do justice to complex data sets. We knew that it would be challenging to come up with a compelling design that meets these needs and also works across formats, but working with renowned editorial designer Mark Porter, we listened, we experimented and we have now acted.

One of the first things you might notice is that the *Nature* logo has changed – this will be the 11th iteration. It's a fresh take on the nature-with-a-small-n that we've used for the past half-century. But it's not just the logo that is new – all our text is now in a custom typeface called Harding, named in memory of Anita Harding, an inspirational professor at London's Institute of Neurology who made important contributions to neurogenetics before her death at the age of just 42.

Working with designers and typographers at Commercial Type, we spent months crafting the typeface to integrate it into *Nature*'s overall design language, inspired by the mid-century Swiss modernist school of rational design. This design school – sometimes called the internationalist school – emerged in response to nationalist design trends before and during the Second World War. It promoted the idea that graphic design should be based on a mathematical grid, allowing designers to arrange type and images with a semblance of order, as *Nature*'s creative director Kelly Krause explains in this issue (see page 476).

The result is a printed journal with text that should be easier to read. We have also adjusted some of the organization and labelling to help readers navigate between sections. From now on, all our research content will also be published in the 'Article' format; the shorter, 'Letter' format has been retired. This will give all the research we publish equal prominence and adds to the extended-data section we created in 2013 to integrate supporting data sets into online papers.

> **"All our text is now in a custom typeface called Harding."**

We have also introduced a new back-page article, called 'Where I work', which profiles researchers, and those connected to research, in the places where they study, work and think. Through a combination of striking photography and first-person narrative, our goal is to provide a glimpse into the lives of people of all ages from around the world. Fans of our Futures articles should not mourn: the journal's science-fiction series continues online.

The redesign process is not over, and you can expect to see more digital changes over the coming year, along with new print and digital design principles for all Nature-branded journals.

Nothing is more important to *Nature* than communicating science with authenticity, accuracy and clarity. We hope the new design does this with a dash of style and with imagination, too. Please tell us what you think. As always, we would welcome your ideas and suggestions for further improvements.

# Precarious supremacy

**Quantum computing will suffer if claims of supremacy are overhyped.**

Researchers led by Google's AI Quantum team have demonstrated 'quantum supremacy' by creating a chip that performed a computational task faster than a classical computer. As we report on page 461, an achievement that the researchers say would have taken the world's fastest supercomputer 10,000 years was completed in under 3 minutes (F. Arute *et al. Nature* **574**, 505–510; 2019).

As the world digests this achievement – including the claim that some quantum computational tasks are beyond supercomputers – it is too early to say whether supremacy represents a new dawn for information technology. It could be that we are looking at quantum computing's Kitty Hawk moment – a reference to the many decades between the Wright brothers' first flight at Kitty Hawk in North Carolina, in 1903 and the advent of the jet age (page 487). At the very least, quantum computers as a routine part of life are likely to be decades or more into the future.

Still, this achievement in science and engineering should certainly not be underestimated. Research teams around the world have been working intensely to unleash the processing power of quantum phenomena: these include superposition, in which particles seem to have multiple states until they are observed; and entanglement, which

nature

describes how the properties of quantum systems can be tied together. If these behaviours can be more precisely controlled, they would generate exponential gains in processing power for certain tasks compared with today's supercomputers. And that is what the team at Google has achieved.

Its chip, known as Sycamore, comprises just 53 individually controllable superconducting quantum bits (qubits), the basic building blocks of quantum computers. The team chose to calculate the outputs of a random quantum circuit – rather like a quantum random number generator. This is not an easy problem, and the Summit supercomputer at Oak Ridge National Laboratory in Tennessee, the world's most powerful machine in its class, would have taken 10 millennia to complete it, the researchers say. Sycamore needed only 200 seconds.

Summit can call on more than 9,000 of the most powerful central processing units (8 billion transistors in each) and nearly 28,000 graphics processors (21 billion transistors each). With such raw computing power outgunned by just 53 qubits, it's understandable that quantum computers are generating such excitement and optimism.

But this demonstration of quantum supremacy is extremely limited. There's a vast gap to be bridged before quantum computers can do more meaningful things – such as simulating the properties of materials or chemical reactions, or accelerating drug discovery.

For one thing, quantum computers are highly sensitive to environmental noise – including everyday phenomena such as temperature changes and electromagnetic fields. And researchers are a long way from being able to design out these and other obstacles.

Instead of proceeding with caution, a quantum gold rush is under way, with investors joining governments and companies to pour large sums of money into developing quantum technologies. Unrealistic expectations are being fuelled that powerful general-purpose quantum computers could soon be on the horizon. Such misguided optimism could be dangerous for the future of this still-fledgling field.

Such a landscape has created a flourishing network of quantum technologists, but those providing the funding will eventually seek a return on investment. There are already concerns that some firms are over-promising, which is why over-hyping this landmark demonstration could raise expectations further. Researchers fear that, if quantum computers fail to deliver anything useful soon, a 'quantum winter' could descend in which research progress slows, investment stalls and disillusion sets in.

The powerful processors that underpin today's devices such as smartphones were developed from decades of sustained investment – often public investment – in research. Quantum processors will similarly require what innovation economists call 'patient capital'.

Too often in the history of science and technology, expectations are raised, only for reality to get in the way. Quantum computers are still near the start of a long and unpredictable journey. As they encounter challenges and costs start to mount, researchers must know that they can reach their destination.

> **"Too often in the history of science, expectations are raised, only for reality to get in the way."**

# Young universities show leadership

**Thriving new institutions can share lessons in building research and publishing capacity.**

This week, an analysis from Nature Research's Nature Index team (see supplement, page S53) looks at the contribution of 'young' universities to research publishing in the natural sciences. Young universities – those aged 50 years or less – are quickly establishing a reputation in teaching and research. However, in Africa, more needs to be done to build their capacity.

The analysis looked at the contributions of authors from 100 young universities in 2018 to 82 journals in the natural sciences. The journals were chosen by an independent panel of researchers, and span the life sciences, physical sciences, chemistry, and Earth and environmental sciences. Author contribution was recorded in several ways, including the total number of articles published by an institution's affiliated researchers, as well as the share of each institution's contribution to those articles.

In most assessments of research-intensive universities, those in the United States and Europe tend to dominate. But among the leading 100 younger universities, there is much more of an east–west mix, spread across China (11 universities), Germany (11), India (10), Australia (9), South Korea (8) and the United States (8).

Authors from the University of the Chinese Academy of Sciences in Beijing are by far the most prolific, contributing 1,816 articles to the listed journals. That is on a par with the number of articles from older institutions in the United States, Europe and Japan, and substantially ahead of second-placed Nanyang Technological University in Singapore (569 articles).

The absence of institutions from Africa in the analysis is partly because many authors there publish in journals that the index does not capture, including in fields such as agriculture, water resources, primary health care and education. But a comparative lack of financial resources for researchers in the natural sciences is also a factor.

In the spirit of south–south collaboration, universities recognized for their publishing in the natural sciences have an opportunity to support those in need of a boost. Many of the young universities assessed in the index are in countries that, even one generation ago, were at an earlier stage of development. They will have valuable lessons to pass on in building research and publishing capacity.

China's fast-expanding universities are already doing this through the Belt and Road Initiative. Rising institutions in other countries, too, will find mutual benefits by sharing experiences and working with research partners in Africa and elsewhere in the global south.

# World view

By Carlos A. Nobre

# To save Brazil's rainforest, boost its science

**My country's government is squeezing research and pushing the rainforest to the brink.**

This August, the skies outside my office in São Paulo, Brazil, were filled with smoke from the fires in the Amazon rainforest. I recalled the forest that I saw in the 1970s as a teenager on family holidays — its beauty, powerful rivers, Indigenous peoples and continuous rains — and thought of how much my country (and the world) could lose. Bishops throughout the Amazon region gathered in the Vatican this month to pray and strategize on behalf of "integral ecology, the cry of the Earth and the poor". Reversing the situation in Brazil is essential for the "good living" the synod seeks.

The smoke has disappeared from my city, for now, but Brazil's rainforest has never been in greater peril. Nor has its science — so badly needed to buttress sustainable agribusiness and invent an economy centred on an intact forest. Last month, thousands of graduate students learnt their scholarships would not be renewed.

Since he came into power in January, Brazil's President Jair Bolsonaro has relaxed the enforcement of laws that prohibit most of the clearing and burning of the Amazon. Analyses of satellite imagery show that this year's dry season brought nearly twice as many fires as last year's, and that the flames were bright on satellite images — as expected from burning a large amount of biomass from recently clearcut forest. Initial estimates indicate that more than 90% of these fires were illegal. Rather than face such alarming data, Bolsonaro fired the director of the agency that monitors deforestation.

These are huge setbacks. From 2005 to 2014, Brazil reduced its annual rate of deforestation by about 75%. Over the same timeframe, the value of agricultural production increased by about 200%. Science and technology fostered this progress. Satellite-based monitoring developed by Brazil's National Institute of Space Research provided daily alerts of deforestation. This facilitated effective law enforcement and incentive programmes for protecting the forest.

My country's example had served as inspiration for others. Without Brazil as a model, I expect deforestation will accelerate across the Amazon. It is already increasing in Colombia, Peru and Bolivia. Twelve years ago, my colleagues and I calculated that, if the rainforest's area shrunk by 40% of its expanse in the 1970s, it could not grow back — and as much as 70% of the original forest could transform to drier, hotter savanna (G. Sampaio *et al. Geophys. Res. Lett.* **34**, L17709; 2007). With rising global temperatures, deforestation, fires and concomitant dryness, that margin has shrunk.

Across most of the Amazon basin, the dry season is already several weeks longer, particularly over deforested areas; these return less moisture to the atmosphere than the rainforest does overall. If deforestation continues, the rainforest could collapse. Many livelihoods will be impossible to maintain. Less rain will fall, temperatures will rise and tens of thousands of species will be lost along with the forest's power to absorb as much as 5% of the world's carbon emissions.

> "The smoke has disappeared from my city, for now, but Brazil's rainforest has never been in greater peril.

Deforestation to increase agricultural lands is no longer necessary. The Brazilian Amazon has an estimated 17 million hectares of degraded and non-productive lands that could be restored and used for sustainable agriculture, including new forest products. Land already in production has the capacity to raise yields severalfold. Innovative technologies and smart management could deliver a 'bio-economy' based on the sustainable extraction of materials for goods, ranging from pharmaceuticals to foods (açaí is the most famous example), cosmetics and other materials. With effective monitoring and enforcement, these 'bio-industries' can boost the economy, respect social rights and traditional peoples and protect the Amazon's ecosystems.

All this will be impossible if Brazil's scientific capacity withers. In 1998, fewer than 4,000 PhD students graduated. Last year, there were more than 22,000. Government funds for science grew steadily from the 1980s, but declined under the economic recession of 2015. Unlike countries such as South Korea, China and Germany that invested more in science to build resilience to economic turbulence, Brazilian science has faced severe budget cuts year after year.

Now, instead of watching younger compatriots build our scientific establishment, I see them all but forced to leave. The government is also changing the direction of research. It aims to replace the forest with livestock farming, monoculture crops, mining operations and huge hydropower plants.

For much of my career, I have worked to reconcile apparently opposing views of land use: some people advocate setting aside large tracts for conservation, and others champion 'resource-intensive development' based on agriculture, livestock, energy and mining. I feel we must instead focus on building a different, sustainable paradigm in which the forest contributes to well-being. Beyond the progress that has already happened in slowing deforestation and boosting sustainable agribusiness, I see much promise in an initiative called Amazonia 4.0, after the Fourth Industrial Revolution of biotechnologies, digital technologies and material science. To realize that promise, we will need scientists and engineers more than ever.

I fear both Brazil's science and the Amazon rainforest are approaching a tipping point — from which recovery is probably impossible. To avoid it, scientists in and outside Brazil should protest vigorously against the anti-science movement and speak clearly to society about how important science and the Amazon are for human well-being and the sustainability of the planet.

**Carlos A. Nobre** is a climate scientist at the University of São Paulo's Institute for Advanced Studies in Brazil and former president of the Brazilian Federal Agency for Support and Evaluation of Graduate Education. e-mail: cnobre.res@gmail.com

CAPES, BRAZIL

# News in brief

## RESEARCH CONCERNS MOUNT AS BREXIT DRAMA UNFOLDS

Brexit uncertainty is paralysing UK politics. It is also taking a toll on the nation's science by making researchers unsure about their future role in European research, according to the Royal Society in London.

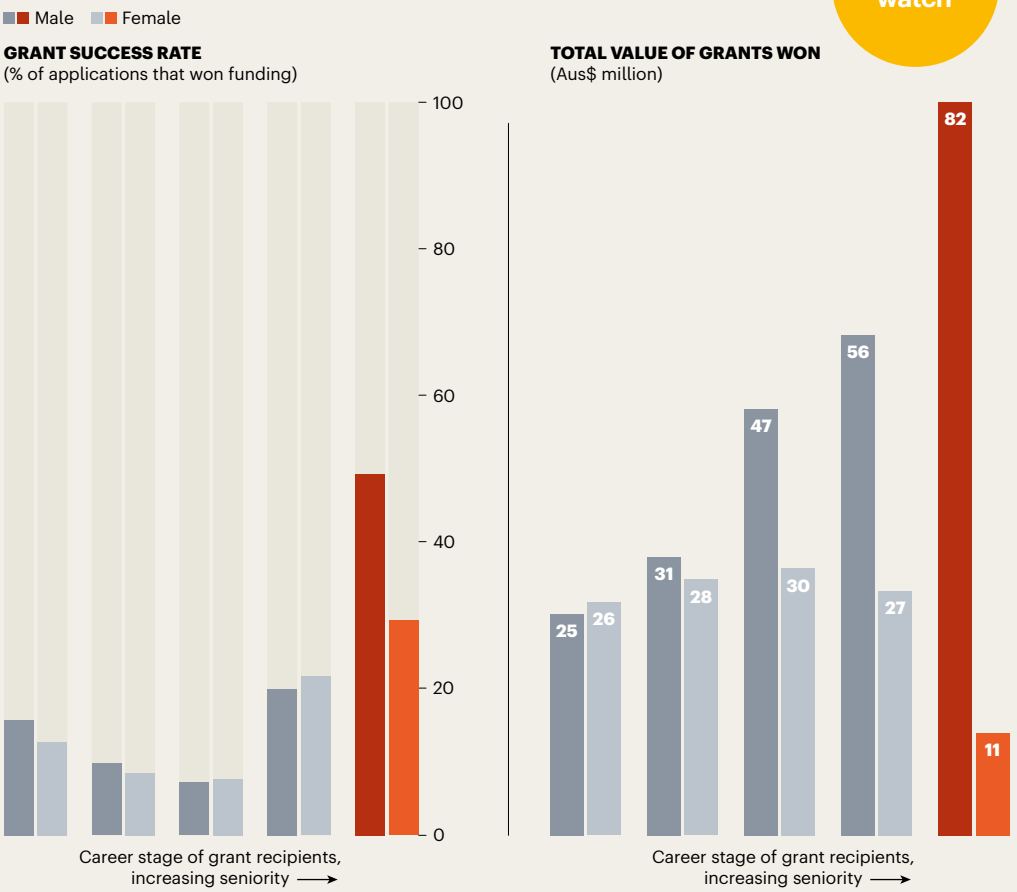
The society says that the United Kingdom's annual share of research funding from the European Union's flagship Horizon 2020 programme fell by almost one-third between 2015 and 2018. This is because UK applications for Horizon 2020 grants dropped by 39% owing to a lack of confidence over the country's future participation in European research, the society says in a report released on 16 October.

As a result, Horizon 2020 funding for UK science dropped by around €500 million (US$560 million), said Royal Society president Venki Ramakrishnan in a statement. He added that there had also been a large drop in the number of leading researchers who want to come to the United Kingdom. "People do not want to gamble with their careers, when they have no sense of whether the UK will be willing and able to maintain its global scientific leadership."

The report shows that last year, the number of non-UK scientists coming to British institutions through the prestigious Marie Skłodowska Curie fellowship scheme, which is part of Horizon 2020, was 35% lower than in 2015.

As *Nature* went to press, the UK government, which is pushing to leave the EU on 31 October, had agreed a deal on the terms of its withdrawal, but still lacked approval from Parliament to proceed.

Trend watch




SOURCE: AUSTRALIAN NATIONAL HEALTH AND MEDICAL RESEARCH COUNCIL

## Women sidelined in medical science funding

Female scientists in Australia were less likely than their male counterparts to win a major new type of medical-research grant, despite an overhaul that was supposed to address gender inequity in the country's science funding. The imbalance occurred in the National Health and Medical Research Council (NHMRC) 'investigator grants', which were awarded for the first time this August. It was particularly severe at the senior-leadership level. Only 29.4% of senior women (5 of 17) who applied were successful, compared with 49.3% of men (37 of 75). "It's a poor message," says Marguerite Evans-Galea, a co-founder of the non-profit association Women in STEMM Australia. Success rates, which were released on the NHMRC's website, were more closely matched at the early- and mid-career stages, but were higher overall for men than for women (14.9% versus 11.3%). Men also received more money overall, partly because they won more grants than women. An NHMRC spokesperson says that extra funding was allocated to several female-led applications that weren't earmarked to receive money, which reduced the gender difference.

## CANADIAN SCIENCE TAKES BACK SEAT IN ELECTION

Canada's Liberal party, led by Justin Trudeau, has won the most seats in the country's general election but not an overall majority in the House of Commons, according to projections available as *Nature* went to press. What the result means for research is unclear.

In the lead-up to the election on 21 October, as the Liberals were running neck and neck with the Conservative Party, researchers had worried that government support would fall by the wayside regardless of which party won.

With the exception of climate change — one of the top issues for voters in recent polls — research was largely absent from the election campaign.

That contrasts with the general election in 2015, when the Trudeau-led Liberal Party campaigned on a promise to reverse policies by the previous government that were widely seen as anti-science — and won.

Since then, the Liberal government has boosted research funding, freed government researchers to speak to the public without first getting permission from the administration, and raised the profile of environmental concerns such as climate change and ocean conservation.

But many researchers felt that the government had begun to rest on its laurels when it came to science. "There is some concern that the government feels like they're done. They've checked the box and they're moving on," says Katie Gibbs, executive director of the campaign group Evidence for Democracy in Ottawa.

The Liberals are now expected to form a coalition government.

L TO R: NASA; ANNE WESTON, EM STP, THE FRANCIS CRICK INSTITUTE/SPL; J. C. WENGA/ANADOLU/GETTY

## US PAIR COMPLETES HISTORIC ALL-FEMALE SPACEWALK

NASA astronauts Christina Koch and Jessica Meir performed the first all-female spacewalk on 18 October, to repair a faulty battery unit on the International Space Station. The roughly seven-hour spacewalk was the fourth for Koch (above right), an electrical engineer who is on track to set a record for the longest single spaceflight by a woman; if all goes to plan, she will spend 328 days in space before returning to Earth in February. Meir (above left), a biologist, had never before attempted a spacewalk.

"This is really just us doing our jobs," Meir said during the walk, which NASA broadcasted live on the Internet.

During the event, the astronauts received a call from US President Donald Trump. "The job that you do is incredible," he told Koch and Meir. "I'm thrilled to be speaking with two brave American astronauts making history."

The two US astronauts are the 14th and 15th women to walk in space. Russian cosmonaut Svetlana Savitskaya was the first, in 1984, followed by 14 Americans.

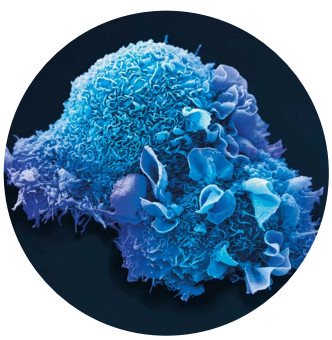
## DETECT CANCER WHEN IT'S SMALL AND TREATABLE

Catching cancer early is the focus of a new transatlantic research collaboration.

The International Alliance for Cancer Early Detection, announced on 21 October, will receive up to £40 million (US$52 million) over five years from the charity Cancer Research UK, with the possibility of an additional £15 million from Stanford University in California and the Oregon Health & Science University Knight Cancer Institute in Portland.

The collaborators hope to take advantage of recent advances in cancer genetics and imaging. Databases are swelling with tumour DNA sequences, and researchers have begun to turn their sights to sequencing precancerous growths in an effort to learn which mutations tip some of them over into malignancy. Clinicians can now detect ever-smaller tumours, and metabolic changes that can be hallmarks of cancer, without surgery or removing tissue.

Early detection could improve cancer treatment: five-year survival rates for six types of cancer are more than three times higher when the cancer is diagnosed at its earliest stage, compared with survival if the cancer is caught only after it has become more advanced and has started to spread to other locations in the body.



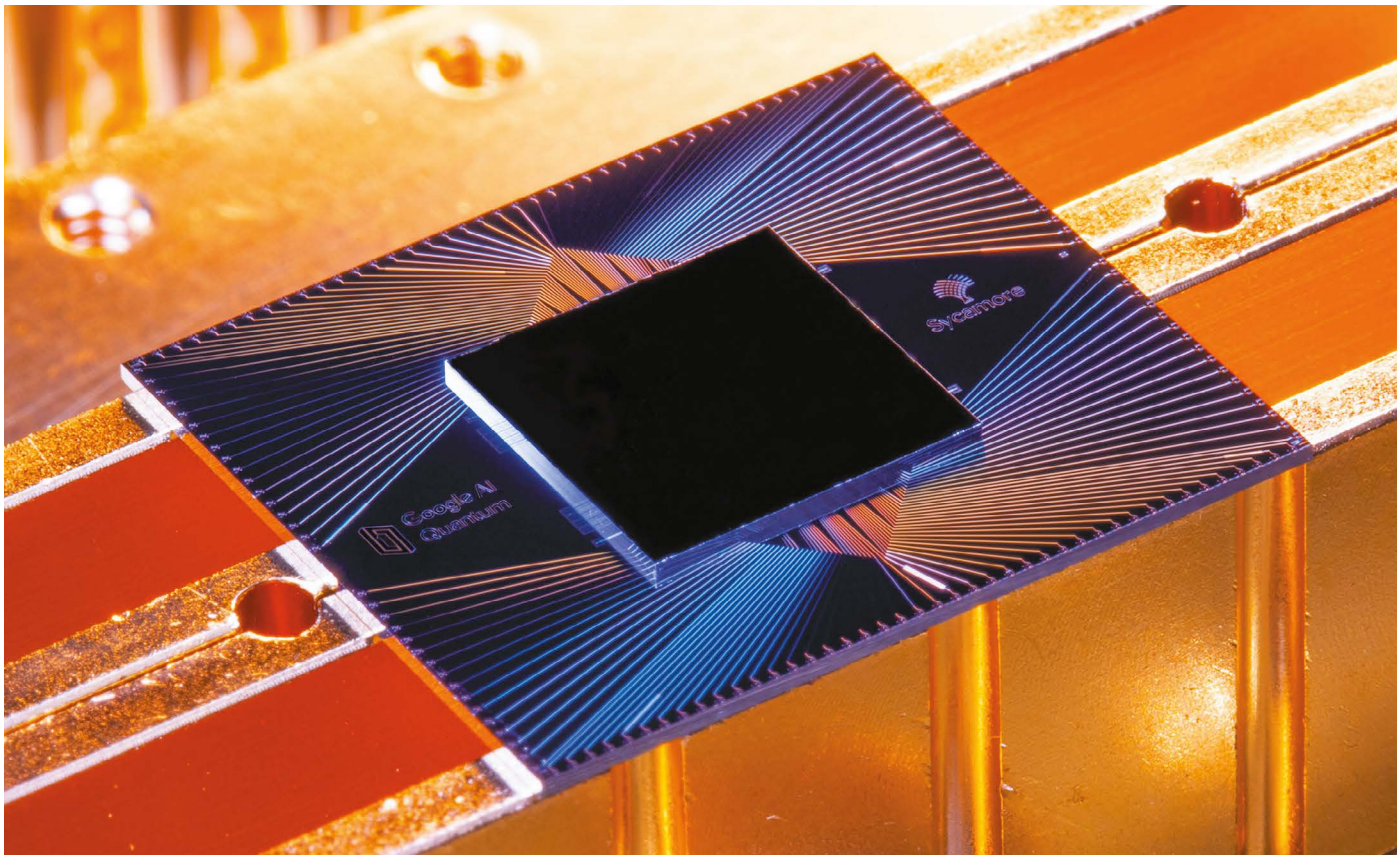
## EBOLA OUTBREAK IN AFRICA SLOWS DOWN

The Ebola outbreak in the eastern Democratic Republic of the Congo (DRC) is finally waning, the World Health Organization (WHO) said on 18 October. Fifty people were diagnosed with Ebola in the DRC between 25 September and 15 October, the WHO said. At the outbreak's peak in April, roughly 300 new infections were reported in three weeks. Almost 3,250 people have been infected since the outbreak began in August 2018, and more than 2,150 have died.

The drop in infections is not a reason to relax efforts to contain the virus, WHO director-general Tedros Adhanom Ghebreyesus told reporters. "We must treat every case as if it is the first since every case has the potential to spark a new outbreak," he said.

There was more good news on 18 October, when the European Medicines Agency (EMA) recommended that the European Commission (EC) approve an Ebola vaccine produced by the pharmaceutical company Merck. About 240,000 people considered to be at risk from Ebola have been given this vaccine during the outbreak, but it is still considered to be an experimental product by regulators worldwide and cannot be marketed. The EC will make a decision within 10 weeks on whether to approve the vaccine for sale.

# News in focus




ERIK LUCERO

**The Sycamore chip is composed of 54 qubits, each made of superconducting loops.**

## GOOGLE PUBLISHES LANDMARK QUANTUM SUPREMACY CLAIM

The company says that its quantum computer is the first to perform a calculation that would be practically impossible for a classical machine.

**By Elizabeth Gibney**

Scientists at Google say that they have achieved quantum supremacy, a long-awaited milestone in quantum computing. The announcement, published in *Nature* on 23 October, follows a leak of an early version of the paper five weeks ago, which Google did not comment on at the time.

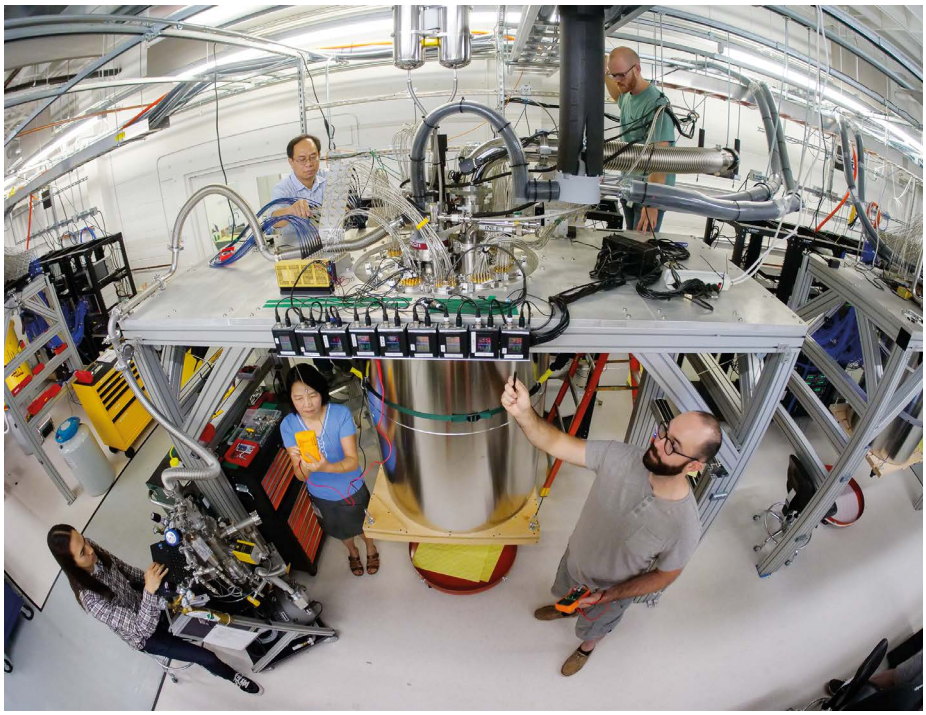
In a world first, a team led by John Martinis, an experimental physicist at the University of California, Santa Barbara, and Google in Mountain View, California, says that its quantum computer carried out a specific calculation that is beyond the practical capabilities of regular, 'classical' machines (F. Arute *et al. Nature* **574**, 505–510; 2019). The same calculation would take even the best classical supercomputer 10,000 years to complete, Google estimates.

Quantum supremacy has long been seen as a milestone because it proves that quantum computers can outperform classical computers, says Martinis. Although the advantage has now been proved only for a very specific case, it shows physicists that quantum mechanics works as expected when harnessed in a complex problem.

"It looks like Google has given us the first experimental evidence that quantum speed-up is achievable in a real-world system," says Michelle Simmons, a quantum physicist at the University of New South Wales in Sydney, Australia.

The feat was first reported in September by the *Financial Times* and other outlets, after an early version of the paper was leaked on the website of NASA, which collaborates with Google on quantum computing, before being quickly taken down. At that time, the company did not confirm that it had written the paper, nor would it comment on the stories.

Although the calculation Google chose — checking the outputs from a quantum random-number generator — has limited practical applications, "the scientific achievement is



ERIK LUCERO

**Google's quantum computer excels at checking the outputs of a random-number generator.**

huge, assuming it stands, and I'm guessing it will", says Scott Aaronson, a theoretical computer scientist at the University of Texas at Austin.

Researchers outside Google are already trying to improve on the classical algorithms used to tackle the problem to bring down the 10,000 year speed-up that the firm calculates. IBM, a rival to Google in building the world's best quantum computers, reported in a preprint on 21 October that the problem could be solved in just 2.5 days using a different classical technique (E. Pednault *et al.* preprint at https://arxiv.org/abs/1910.09534; 2019). That paper has not been peer-reviewed. If IBM is correct, it would reduce Google's feat to demonstrating a quantum 'advantage' — doing a calculation much faster than a classical computer, but not something that is beyond its reach. This would still be a significant landmark, says Simmons. "As far as I'm aware that's the first time that's been demonstrated, so that's definitely a big result."

## Quick solutions

Quantum computers work in a fundamentally different way from classical machines: a classical bit is either a 1 or a 0, but a quantum bit, or qubit, can exist in multiple states at once. When qubits are inextricably linked, physicists can, in theory, exploit the interference between their wave-like quantum states to perform calculations that might otherwise take millions of years. Physicists think that quantum computers might one day run revolutionary algorithms that could, for example, search unwieldy databases or factor large numbers — including, importantly, those used in encryption. But those applications are still decades away. The more qubits are linked, the harder it is to maintain their fragile states while the device is operating. Google's algorithm runs on a quantum chip composed of 54 qubits, each made of superconducting loops. But this is a tiny fraction of the one million qubits that could be needed for a general-purpose machine.

The task Google set for its quantum computer is "a bit of a weird one", says Christopher Monroe, a physicist at the University of Maryland in College Park. Google physicists first crafted the problem in 2016, and it was designed to be extremely difficult for an ordinary computer to solve. The team challenged its computer, known as Sycamore, to describe the likelihood of different outcomes from a quantum version of a random-number generator. They do this by running a circuit that passes 53 qubits through a series of random operations. This generates a 53-digit string of 1s and 0s — with a total of $2^{53}$ possible combinations (only 53 qubits were used because one of Sycamore's 54 was broken). The process is so complex that the outcome is impossible to calculate from first principles, and is therefore effectively random. But owing to interference between qubits, some strings of numbers are more likely to occur than others. This is similar to rolling a loaded die — it still produces a random number, even though some outcomes are more likely than others.

**"The scientific achievement is huge, assuming it stands, and I'm guessing it will."**

Sycamore calculated the probability distribution by sampling the circuit — running it one million times and measuring the observed output strings. The method is similar to rolling the die to reveal its bias. Verifying the solution was a further challenge. To do that, the team compared the results with those from simulations of smaller and simpler versions of the circuits, which were done by classical computers — including the Summit supercomputer at Oak Ridge National Laboratory in Tennessee. Extrapolating from these examples, the Google team estimates that simulating the full circuit would take 10,000 years even on a computer with one million processing units (equivalent to around 100,000 desktop computers). Sycamore took just 3 minutes and 20 seconds.

## Limited applications

Monroe says that Google's achievement might benefit quantum computing by attracting more computer scientists and engineers to the field. But he also warns that the news could create the impression that quantum computers are closer to mainstream practical applications than they really are. "The story on the street is 'they've finally beaten a regular computer: so here we go, two years and we'll have one in our house'," he says.

In reality, Monroe adds, scientists are yet to show that a programmable quantum computer can solve a useful task that cannot be done any other way, such as by calculating the electronic structure of a particular molecule — a fiendish problem that requires modelling multiple quantum interactions. Another important step, says Aaronson, is demonstrating quantum supremacy in an algorithm that uses a process known as error correction — a method to correct for noise-induced errors that would otherwise ruin a calculation. Physicists think this will be essential to getting quantum computers to function at scale. Google is working towards both of these milestones, says Martinis, and will reveal the results of its experiments in the coming months.

Aaronson says that the experiment Google devised to demonstrate quantum supremacy might have practical applications: he has created a protocol to use such a calculation to prove to a user that the bits generated by a quantum random-number generator really are random. This could be useful, for example, in cryptography and some cryptocurrencies, whose security relies on random keys.

Google engineers had to carry out a raft of improvements to their hardware to run the algorithm, including building new electronics to control the quantum circuit and devising a new way to connect qubits, says Martinis. "This is really the basis of how we're going to scale up in the future. We think this basic architecture is the way forward," he says.

GEORGE ROSE/GETTY

**Western Canada's Rocky Mountains are among the geologically complex areas that Earth scientists hope to study in detail.**

# EARTH SCIENTISTS PUSH PLAN TO MAP CANADA'S GEOLOGY

## A fleet of geophysical observatories would probe everything from the inner Earth to the upper atmosphere.

**By Alexandra Witze**

Sometime in mid- to late November, geophysicist David Eaton will head into the forests around Fort St. John, Canada, and help to nestle an array of 15 seismometers into the ground. They will spend their days listening for small earthquakes caused by oil and gas exploration in this part of British Columbia. If Eaton has his way, the seismometers will soon be joined by hundreds more, blanketing Canada as part of an unprecedented quest to probe the nation's geology.

Eaton, of the University of Calgary, is leading a hugely ambitious effort to establish a network of geophysical observatories across Canada. The project aims to study everything from the inner Earth to the upper atmosphere — and to answer questions such as how much Canadians should worry about earthquakes and landslides, and where researchers should explore for lucrative mineral deposits or renewable energy resources.

"We're completely aspirational and ambitious," Eaton says.

It's not clear whether funding agencies will cough up the roughly Can$100 million (US$75 million) that's needed to turn these ambitions into reality. But a wide-ranging group of scientists has come together to advocate for the project, which is known as EON-ROSE (Earth-system Observing Network-Réseau d'Observation du Système Terrestre). Some, including Eaton, are now journeying into the Canadian wilderness to show what EON-ROSE could do.

EON-ROSE was inspired by a massive US geophysics programme called EarthScope that is wrapping up this year. Since 2004, the project has covered the continental United States with a movable network of observatories composed of seismometers and GPS instruments. EarthScope revealed new details about geological hazards — such as the enormous Cascadia fault zone in the Pacific Northwest, where researchers discovered unexpected slow-moving quakes that might offer clues to how often a massive quake is likely to strike the region.

That caught the attention of scientists to the north. "We started talking about, why can't we do this in Canada?" says Roy Hyndman, a geophysicist at the Pacific Geoscience Centre in Sidney, British Columbia, who led many of the early discussions. Like EarthScope, EON-ROSE aims to install geophysical observatories in a grid that would move around the country.

The project's backers want to study Canada's geologically active western mountains, along

with the ancient rocks in its centre that date back roughly four billion years, and unstable regions in the east. Compared with the United States, "Canada covers an even wider portion of Earth history that could be investigated", says Andy Frassetto, a seismologist at the Incorporated Research Institutions for Seismology in Washington DC.

EON-ROSE organizers have begun the project with a series of smaller studies, such as Eaton's, while they seek full funding. Another study took place this summer, when researchers from the Geological Survey of Canada and Geoscience BC descended on Mount Meager, which is in southern British Columbia and is Canada's most recently active big volcano. Their goal was to explore whether its volcanic warmth — which heats groundwater up to 240 °C — could be tapped for geothermal energy.

In July, geologists travelled around the mountain in helicopters to install instruments similar to those envisioned for EON-ROSE. The researchers are crunching the preliminary data now, aiming to see where permeable rocks channel Mount Meager's volcanic heat towards the surface. Future studies in other parts of Canada could help geologists to find new sources of geothermal energy — such as in the remote Arctic, where residents often rely on imported diesel, says Stephen Grasby, a geochemist at the Geological Survey of Canada in Calgary who led the work.

### Hidden treasure

EON-ROSE also aims to identify mineral deposits by looking for geological structures deep below the surface that might underlie lodes of gold or copper. This approach could make it easier to prospect for minerals in the country's northern reaches, where harsh winters and a shortage of roads make it difficult to explore.

"You could spend forever up there wandering around before you discover anything," says Keith Benn, a mineral-exploration consultant in Port Lambton, Canada. "This is the promise of the EON-ROSE approach — when you look at this expansive territory in northern Canada, we can say, 'we can help you decide where to start'."

Benn is working with mining companies to drum up funding for a pilot EON-ROSE study of the ancient rocks of central Canada.

This focus on energy and mineral exploration goes beyond the purely scientific aims of EarthScope. EON-ROSE organizers hope that a broader focus will help them win funding from industry. "To move forward, we must have practical applications that benefit Canada," says Katherine Boggs, a geologist and project leader at Mount Royal University in Calgary.

Ultimately, the scientists hope to get the bulk of their funding from the federal government — although Canada's general election on 21 October could markedly shift the outlook for science funding.

# PRECISE CRISPR TOOL COULD TACKLE HOST OF GENETIC DISEASES

## Greater control could allow many more conditions to be treated with gene editing.

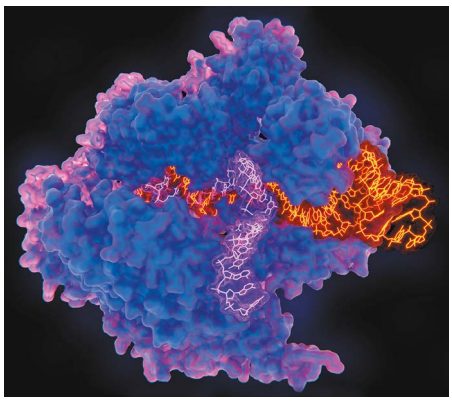
**By Heidi Ledford**

For all the ease with which the wildly popular CRISPR–Cas9 gene-editing tool alters genomes, it's still somewhat clunky and prone to errors and unintended effects. Now, an alternative offers greater control over genome edits — an advance that could be particularly important for developing gene therapies.

The alternative method, called prime editing, improves researchers' chances of getting only the edits they want, instead of a mix of changes that they can't predict. The tool, described in *Nature* (A. V. Anzalone *et al. Nature* http://doi.org/dczp; 2019) on 21 October, also reduces the 'off-target' effects that are a key challenge for some uses of the standard CRISPR–Cas9 system. That could make prime-editing-based gene therapies safer.

The tool also seems capable of making a wider variety of edits, which might one day allow it to be used to treat the many genetic diseases that have so far stymied gene editors. David Liu, a chemical biologist at the Broad Institute of MIT and Harvard in Cambridge, Massachusetts, and lead author of the study, estimates that prime editing might help researchers tackle nearly 90% of the more than 75,000 disease-associated DNA variants listed in ClinVar, a database developed by the US National Institutes of Health.

And the specificity of the changes that prime editing is capable of could make it easier for researchers to develop models of disease, or to study specific gene functions, says Liu.



JUAN GAERTNER/SPL

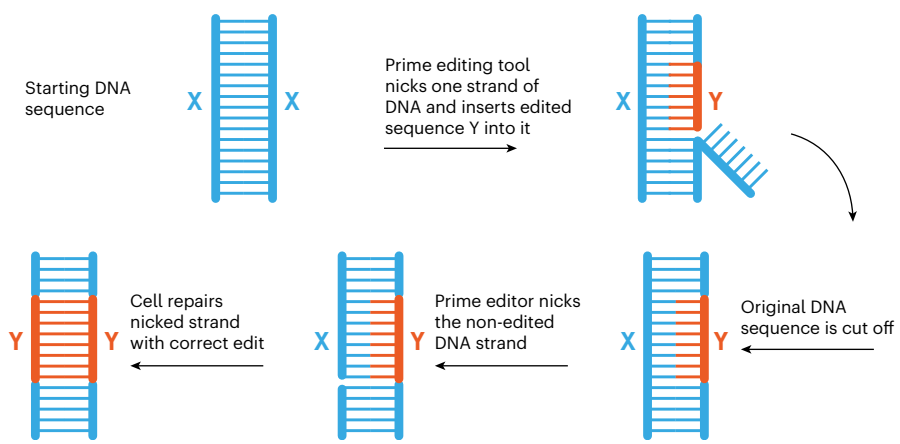
**A new gene-editing tool offers more control than CRISPR–Cas9 (pictured).**

"It's early days, but the initial results look fantastic," says Brittany Adamson, who studies DNA repair and gene editing at Princeton University in New Jersey. "You're going to see a lot of people using it."

Prime editing might not be able to make the very big DNA insertions or deletions that CRISPR–Cas9 is capable of — so it's unlikely to completely replace the well-established editing tool, says molecular biologist Erik Sontheimer at the University of Massachusetts Medical School in Worcester. That's because for prime editing, the change that a researcher wants

### PRECISION EDITOR

Prime editing reduces the number of unintended changes to a genome by inserting the edits researchers want to make into the DNA itself. This contrasts with CRISPR–Cas9, which relies on the cell's repair system to make the changes.



to make is encoded on a strand of RNA. The longer that strand gets, the more likely it is to be damaged by enzymes in the cell.

"Different flavours of genome-editing platforms are still going to be needed for different types of edits," says Sontheimer.

But prime editing seems to be more precise and versatile than other CRISPR alternatives. Those include modified versions of CRISPR–Cas9 that enable researchers to swap out one DNA letter for another, and older tools such as zinc-finger nucleases, which are difficult to tailor to each desired edit.

### Freedom through control

CRISPR–Cas9 and prime editing both work by cutting DNA at a specific point in the genome. CRISPR–Cas9 breaks both strands of DNA at once and then relies on the cell's own repair system to patch the cuts and make the edits. But that repair system is unreliable and can insert or delete DNA letters at the points where the genome was cut. This can lead to an uncontrollable mixture of edits that vary between cells.

Even when researchers include a template to guide the edits, the DNA repair system in most cells is much more likely to make those small, random insertions or deletions than to add a specific sequence to the genome. That makes it difficult for researchers to use CRISPR–Cas9 to overwrite a piece of DNA with a sequence of their choosing.

Prime editing bypasses these problems (see 'Precision editor'). It, too, uses Cas9 to recognize specific DNA sequences, but the prime editor's Cas9 enzyme is modified to nick only one DNA strand. Then, a second enzyme called reverse transcriptase, guided by a strand of RNA, makes the edits at the site of the cut.

The prime-editing enzymes don't have to break both DNA strands at the same time to make changes, freeing researchers from relying on the cell's genome repair system — which they can't control — to make the edits that they want. This means that prime editing could enable the development of treatments for genetic diseases caused by mutations that aren't easily addressed by existing gene-editing tools.

Previously, researchers, including Liu, thought that they would need to develop gene-editing tools specific to each category of change they wanted to make in a genome: insertions, deletions or DNA letter substitutions. And the options were limited when it came to making precise substitutions.

An older technique, called base editing, which is comparable in precision to prime editing, chemically converts one DNA letter directly into another — changing a T to an A or a G to a C — without breaking both DNA strands. That's something CRISPR–Cas9 can't do. Developed by Liu, base editing could be useful for correcting genetic diseases caused by single-letter mutations, including the most common form of sickle-cell anaemia.

But base editing can't help with genetic disorders caused by multi-letter mutations such as Tay–Sachs disease, a usually fatal illness typically caused by the insertion of four DNA letters into the *HEXA* gene. So Liu and his colleagues set out to create a precise gene-editing tool that gave researchers the flexibility and control to make multiple types of edits without having to create bespoke systems.

"It's fantastic," says Sontheimer. "The breadth of the mutations that can be introduced is one of the biggest advances. That's huge."

Liu's team, and others, will now need to carefully evaluate how well the system works in a variety of cells and organisms. "This first study is just the beginning — rather than the end — of a long-standing aspiration in the life sciences to be able to make any DNA change at any position in an organism," says Liu.

# RUSSIAN SCIENTIST EDITS HUMAN EGGS IN EFFORT TO ALTER DEAFNESS GENE

Denis Rebrikov says he does not plan to implant gene-edited embryos until he gets regulatory approval.

**By David Cyranoski**

Russian biologist Denis Rebrikov has started editing genes in human eggs with the goal of repairing a mutation that can cause deafness. The news, detailed in an e-mail he sent to *Nature* on 17 October, is the latest chapter in a saga that kicked off in June, when Rebrikov revealed his controversial intention to create gene-edited babies resistant to HIV using the popular CRISPR tool. So far, only one person has claimed to have created a baby from a gene-edited embryo — the Chinese scientist He Jiankui, in November 2018.

Rebrikov's e-mail (see Q&A on page 466) follows a September report in the Russian magazine *N+1,* in which he said a couple who both have a genetic mutation that impairs their hearing had started procedures to collect eggs that would be used in an attempt to create a gene-edited baby. The eggs that Rebrikov has edited so far are from women without the genetic mutation. He says the goal of those experiments is to learn how to allow couples with the mutation to have a child with unaffected hearing.

He also wants to better understand potentially harmful 'off-target' mutations, which are a known challenge of using the CRISPR–Cas9 system to edit embryos.

Rebrikov says he does not plan to use the tool to create such a baby yet — and that his previously reported plan to apply this month for permission to implant gene-edited embryos in women has been pushed back.

Instead, he says, he will soon publish the results of his egg experiments, which also involved testing CRISPR's ability to repair the gene linked to deafness, called *GJB2*, in body cells taken from people with the mutation. People with two mutated copies of *GJB2* cannot hear well without interventions such as hearing aids or cochlear implants. Rebrikov says that these results will lay the groundwork for implanting an edited embryo.

Rebrikov adds that he has permission from a local review board to do his research, but that this does not allow transfer of gene-edited eggs into the womb and subsequent pregnancy.

Apart from the couple who agreed to start undergoing egg collection, he is in discussion with four other couples in which both would-be parents have two mutated *GJB2* genes, he says.

Rebrikov also provided further information about the couple who agreed to the procedures. In September, *N+1* reported that the couple hadn't signed a consent form and had backed away from the idea of creating a gene-edited child, citing personal reasons.

**"I will definitely not transfer an edited embryo without the permission of the regulator."**

But Rebrikov now says that this is only a temporary hurdle. He notes that the woman who donated the eggs has taken a one-month pause while she gets a cochlear implant.

Rebrikov also emphasized that he will not move forwards without approval from the Ministry of Health of the Russian Federation. "I will definitely not transfer an edited embryo without the permission of the regulator."

That might not come any time soon. Earlier this month, the ministry released a statement saying that production of gene-edited babies is

Denis Rebrikov plans to publish his experiments to repair genes in human eggs soon.

ANDREY RUDAKOV/BLOOMBERG/GETTY

premature. Rebrikov says "it is hard to predict" when he will get permission, but that it will be after all the necessary safety checks.

Rebrikov shot to fame in June when he told *Nature* of his plans to make HIV-resistant babies. The news shocked international researchers, who feared that he was following in the footsteps of He Jiankui.

Those plans involve using CRISPR to disrupt the same gene that He did — *CCR5*. The protein made by the *CCR5* gene allows HIV to enter cells, and people with a mutated copy of this gene

**"The project is clearly unethical and damages the credibility of a technology intended to help, not harm."**

are much less likely to get the virus. But many scientists say that the benefits — possible resistance to HIV — are not worth the unknown risks of gene editing, because there are other ways to prevent HIV passing from parent to child.

Rebrikov says he started looking for women with HIV who wanted to have a baby and who have responded poorly to HIV drugs. He argues that such people might be good candidates for the procedure because they have an elevated risk of passing the virus to their children, although many scientists think that any attempt to use gene editing in embryos to modify *CCR5* is misguided. In his latest e-mail, Rebrikov told *Nature* that he is still looking for suitable women. "But there are very few of them," he says.

In the meantime, Rebrikov has taken on the project to repair the *GJB2* gene in human embryos. Some scientists also question the benefits of this procedure because hearing loss is not a fatal condition. "The project is recklessly opportunistic, clearly unethical and damages the credibility of a technology that is intended to help, not harm," says Jennifer Doudna, a pioneer of the CRISPR tool at the University of California, Berkeley.

In the wake of He's explosive revelation, the World Health Organization (WHO) tasked a committee with developing an international framework to govern the clinical use of gene editing. In August, the WHO committee also launched an international registry of clinical research using gene editing in humans to oversee this practice. An international commission created by the US National Academy of Sciences, the US National Academy of Medicine and the United Kingdom's Royal Society is also preparing a framework to guide clinical research in germline gene editing. This is expected to be released by mid-2020. The commission will hold a public meeting on 14–15 November to gather ideas.

Rebrikov said last month that he wants to follow regulations that have been internationally agreed on when moving gene editing to the clinic, according to the Bloomberg news agency. But he also expressed frustration that none exists yet.

Robin Lovell-Badge, a developmental biologist at the Francis Crick Institute in London and a member of the WHO committee, says that Rebrikov should wait until such a framework has been agreed, which will take time. "This is not a simple matter, and it is ridiculous to think that we can come up with global solutions to regulation in a very complex scientific and potentially clinical area in a few months."

Q&A

## Denis Rebrikov

**Below are edited versions of the questions that *Nature* sent to Rebrikov, and his answers.**

**Some scientists and bioethicists say that, because deafness is not a life-threatening condition, it should not be the target of a risky treatment like this.**
Any new drug carries certain risks. The deafness model is the most appropriate for applying genomic editing at the zygote [newly fertilized egg] stage.

**In particular, scientists worry about off-target mutations — which are potentially dangerous and could be introduced away from the intended edit.**
Of course we worried about those. We have a long and reasonable algorithm for checking off-target activity. I'd like to discuss the algorithm for checking the efficiency and safety of the technology, rather than the method's prematureness.

**Some also warn that because the CRISPR repair mechanism is inefficient, there is a high likelihood of producing children with mosaicism — a mix of edited and unedited cells. Are you worried about this?**
Yes. Unfortunately, due to the impossibility of a complete analysis of the embryo — we only look at a biopsy of five to seven cells — we will never be completely sure of the absence of mosaicism in transferred embryos. But statistically (in experiments), it is possible to show either the permissible percentage of mosaicism or its absence.

**The Russian health ministry said earlier this month that it follows the position of the World Health Organization committee: it is too early to do such experiments. Will you apply anyway?**
What does it mean, too soon? Lenin said, "yesterday was too early, tomorrow it will be too late."

**Those working on international frameworks to guide the clinical application of human-embryo editing have suggested that, until they are done, clinical research should slow down.**
Are you serious? Where did you see the researcher willing to slow down?

This rat-sized *Liaoconodon hui* is one of many fossils from northern China that are sharpening the picture of how mammal traits evolved.

# THE MAKING OF MAMMALS

An explosion of fossil finds is allowing our early mammal ancestors to leap out of the shadow of the dinosaurs. **By John Pickrell**

Night is falling in the early Jurassic 185 million years ago, and the *Kayentatherium* is tending to her newly hatched brood. Heavy rains pummel the bank above her den as she looks over her dozens of tiny young. She is about the size of a large cat and could easily pass for a mammal, but her large jawbone, characteristic teeth and lack of external ears give her away: she is a cynodont, a member of the group from which mammals evolved. At some point without warning, the sodden bank collapses, entombing the hatchlings and their mother in mud.

There they remained until the summer of 2000, when a fossil-hunting crew led by Timothy Rowe at the University of Texas at Austin chanced upon their scattered bones among rocks of the Kayenta Formation in northern Arizona.

That initial encounter with the fossils did little to impress the palaeontologists. They dug up the block and shipped it back to the laboratory for safekeeping. It wasn't until nine years later that a specialist preparing the fossil for study noticed something startling: embedded in the block were tiny teeth, and jawbones just 1 centimetre in length. "Immediately they stopped the preparation and thought about ways of non-destructively examining the babies," says Eva Hoffman, at Texas with Rowe at the time and now a palaeontologist at the American Museum of Natural History in New York City. Instead of breaking into the rock, Hoffman and Rowe digitally extracted the bones with a microcomputed tomography (microCT) scanner, which uses X-rays to create fine-grained 3D images.

What they found inside the rock were the first known babies of mammals or their relatives from the Jurassic — and not just one, but 38 of them, placing this among the most significant discoveries related to mammal origins made in the past decade[1]. *Kayentatherium* is at the cusp of mammalhood — and researchers say that it provides crucial insights into which traits define mammals and which were present in their earlier relatives.

*Kayentatherium*'s skeleton is mammal-like in many ways, but the fossil suggested that it still reproduced very much like a reptile, giving birth to large litters of small-brained offspring. By contrast, "mammal moms invest a lot in a smaller number of babies, each of which has a better chance of surviving", says Hoffman. Mammal babies spend longer under their parents' care, developing relatively large brains, whereas these fossil hatchlings had well-developed bones and teeth, hinting that they could fend for themselves and were not nourished by milk, as all mammals are today.

The find is among a mass of discoveries in the past 10–20 years that are illuminating milestones in mammalian evolution. Although major finds are emerging all over the world, the largest number are coming out of China;

J. MENG *ET AL. NATURE* **472**, 181–185; 2011

together, they have overturned the now dated belief that dinosaur-era mammals were small, unremarkable insectivores, eking out a life in the shadows of the giant reptiles.

The fossils have revealed that early mammals were ecologically diverse and experimenting in gliding, swimming, burrowing and climbing. The discoveries are also starting to reveal the evolutionary origins of many of the key traits of mammals – such as lactation, large brains and superbly keen senses.

"The explosion of early-mammal discoveries, particularly from China, over the last two decades has been eye-opening, mind-numbing and absolutely dazzling," says David Krause, a vertebrate palaeontologist at the Denver Museum of Nature and Science in Colorado.

This avalanche of discovery is also stirring up debate: some researchers disagree over which fossil groups are true mammals and the shape of the mammal family tree. "We want to understand our early history in the language of evolutionary biology, and that's what fires me up," says Zhe-Xi Luo, a palaeontologist at the University of Chicago in Illinois. "That's why this entire field is so interesting, because the fossil record is getting better and better, and we are starting to really tackle some of these questions."

## Out of the shadows

In 1824, at the Geological Society of London, naturalist William Buckland presented bones from one of the first known dinosaurs, *Megalosaurus*. At the same talk, he revealed tiny mammalian jaws that had been found in the same fossil deposit. Their presence suggested that mammals had a very deep history, but as would happen repeatedly, the dinosaur discoveries completely overshadowed the mammal ones.

The slow trickle of mammal finds from around the world continued for 150 years. Then in 1997, researchers described the first ancient mammal from the fossil-rich rocks of Liaoning in northeastern China[2], and the floodgates opened. Since then, 50 or more near-complete and "beautiful specimens" have been found there, according to Jin Meng, a palaeontologist at the American Museum of Natural History. Like the dinosaur fossils, they are dug up by local farmers and sold on to museums.

But the dinosaurs continued to get the vast majority of the attention, says palaeontologist Steve Brusatte at the University of Edinburgh, UK. "It's only that very recently, through the work of Luo, Meng and others, that the mammals are getting their due."

Most of China's mammal fossils were formed when volcanoes buried the animals in ash – and they are exquisitely detailed. Typical mammal fossils from the Mesozoic era (252 million to 66 million years ago) are little more than teeth and jaw fragments, but Chinese specimens often have entire skeletons, with fur, skin and internal organs. "We have a lot of detail to answer scientific questions," says Meng, He is interested in understanding the evolution of the mammalian ear, for instance.

The finds overturned previous dogma. "We used to say that during the time of dinosaurs, mammals were totally unspectacular. That they were just these little mousey things scampering around in the shadows," says Brusatte. But these animals "were undergoing their own evolutionary explosion", he says.

Mammals first appeared at least 178 million years ago, and scampered amid the dinosaurs until the majority of those beasts, with the exception of the birds, were wiped out 66 million years ago. But mammals didn't have to wait for that extinction to diversify into many forms and species. "These new discoveries document a huge, hitherto-undreamed-of ecological diversity," says Richard Cifelli, a palaeontologist at the University of Oklahoma in Norman.

Among the first innovations that researchers began to find in fossil form were those to do with locomotion. In 2006, Meng's team reported the first gliding mammal[3], 164-million-year-old *Volaticotherium*, which had wing membranes made of furry skin, like today's flying squirrels. In 2017, Luo's team added *Vilevolodon* and *Maiopatagaium*[4,5], which lived at around the same time and belonged to a group called the haramiyids. These animals swooped between the trees alongside some of the first flying dinosaurs, taking advantage of previously unexploitable food resources.

> "THESE NEW DISCOVERIES DOCUMENT A HUGE, HITHERTO-UNDREAMED-OF ECOLOGICAL DIVERSITY."

Researchers found other specializations that they assumed had evolved only later: *Agilodocodon* could climb trees and gnawed into bark to feast on sap[6]; the platypus-sized river-dweller *Castorocauda* had webbed feet and a beaver-like tail for swimming[7]; and *Docofossor* had paws and claws for digging, and looked like a modern mole[8].

These mammals had also adapted to a multitude of diets, much more diverse than previously assumed. In 2014, Krause described the groundhog-like *Vintana* from Madagascar[9], a herbivore that perhaps fed on roots and seeds. And the wolverine-sized carnivore *Repenomamus*, which Meng's team reported in 2005, had baby dinosaur bones in its stomach[10]. Many of these new-found fossil mammals belong to long-extinct subgroups, says Meng. In contrast to the panoply that existed in the Mesozoic, mammals today come in just three varieties: placentals, which make up the majority of species and include humans; marsupials, such as kangaroos and koalas, in which gestation in the womb is brief and development continues in a pouch; and the egg-laying monotremes, represented only by the platypus and several echidnas. "But in geological history, there were many other groups such as multituberculates, triconodonts and haramiyids," says Meng. "Mammals were actually very diverse in the Jurassic."

Some, such as the shrew-like *Juramaia* – described by Luo's team in 2011 and dated to 160 million years ago – are among the earliest placental mammals and therefore have the potential to be our ancestors[11].

And a few dinosaur-era mammals were much bigger than suspected, too. *Repenomamus* was 12–14 kilograms, and the racoon-sized *Vintana* weighed in at 9 kg. "It's exciting that we kind of busted the old myths that early mammals came from a very humble generalized ancestor," says Luo.

The finds are not solely from China. Important fossils are also coming from the United States, Spain, Brazil, Argentina, Madagascar and Mongolia. Some of the most intriguing and oldest fossils – as well as the biggest gaps in our knowledge – relate to the southern continents, where only five genera of Mesozoic mammals and their relatives are known, compared with more than 70 genera from northern latitudes. In the past two decades, Brazil has yielded several Triassic fossils that are more than 200 million years old. Guillermo Rougier, a palaeontologist at the University of Louisville in Kentucky, describes them as "incredible discoveries" that are right on the cusp between mammals and their cynodont ancestors. "These forms really show a very transitional progression from things that are typically non-mammalian, to things that pretty much have all the features of early mammals."

## Mammal must-haves

The latest finds are also offering clues to the evolution of key mammal features. For instance, the keen hearing of mammals is partly down to tiny bones in the middle ear – the malleus, incus and ectotympanic. But in the reptilian ancestors of mammals, these bones were part of the jaw, and were used for chewing instead of hearing. Mammal forerunners, such as shrew-like *Morganucodon* from 205 million years ago, sported a prototype of the mammal arrangement that allowed for both functions[12].

In 2011, Meng reported an intermediary[13]: a 120-million-year-old specimen from China belonging to a group of mammals called eutriconodonts and named *Liaoconodon hui* (see 'Mammal hallmarks'). The rat-sized fossil revealed three middle-ear bones, but they were still attached to the jaw **CONTINUED ON P.472 ▸**

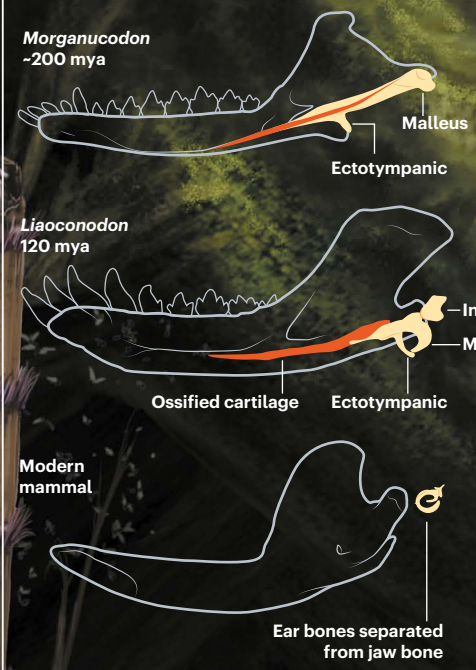
# MAMMAL HALLMARKS

A raft of extraordinary fossil finds is revealing details of how mammals evolved from reptilian forebears more than 178 million years ago. Discoveries in the past two decades show that early mammals were a diverse bunch, with sophisticated skills such as gliding and burrowing that researchers thought evolved only later. Many of the features that define mammals – like suckling milk, exceptional hearing and small litter sizes – had already appeared by the time true mammals were roaming the land, rivers and skies.

By John Pickrell
Illustration by Davide Bonadonna
Design by Wes Fernandes

## WALKING WITH DINOSAURS

Early mammals like this rat-sized species *Liaoconodon hui* coexisted with feathered dinosaurs like *Sinotyrannus* in the temperate ecosystems of the Cretaceous in what is now Liaoning in northern China.

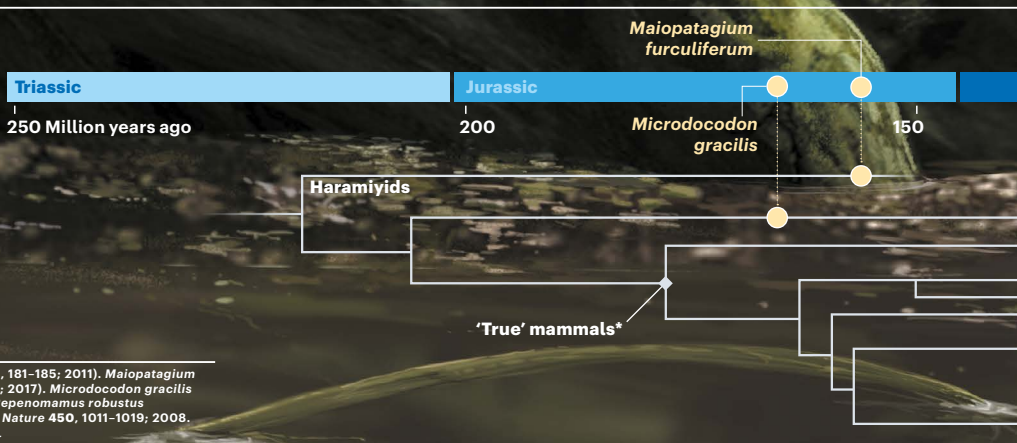



## JAW BECOMES EAR

Mammals have three middle ear bones – the malleus, incus and ectotympanic bone (yellow) – responsible for their fine hearing. In reptiles and the ancestors of mammals, these bones formed part of the jaw and helped to process food. When researchers reported *Liaoconodon hui* in 2011 (main picture), they realized it represented a transitional state: a piece of hardened cartilage in the jaw (orange) supported some of the bones and the eardrum.

## DEBATABLE DYNASTY

Researchers disagree about the shape of the mammal family tree, which species fall in or out of it, and even where the tree begins. One view suggests that the earliest true mammals are 178 million years old; another argues that the oldest mammal dates from 208 million years ago. Many recent fossil finds represent long-extinct groups that branched off before the tree gave rise to the three groups of modern mammals.




*Liaoconodon hui* and evolution of the ear bones (J. Meng *et al. Nature* **472**, 181–185; 2011). *Maiopatagium furculiferum* and gliding mammals (Q.-J. Meng *et al. Nature* **548**, 291–296; 2017). *Microdocodon gracilis* hyoid bone and suckling (C. F. Zhou *et al. Science* **365**, 276–279; 2005). *Repenomamus robustus* (Y. Hu *et al. Nature* **433**, 149–152; 2005). Timeline adapted from Z.-X. Luo *Nature* **450**, 1011–1019; 2008.

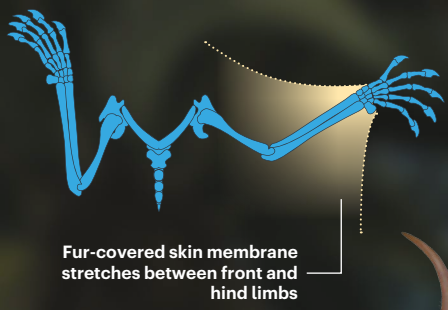
*Researchers do not know the exact timing of when lineages split from each other

# GALLERY OF TALENTS

Many archetypal mammalian features evolved in a short burst early in mammal evolution, including innovations in movement, development and diet.

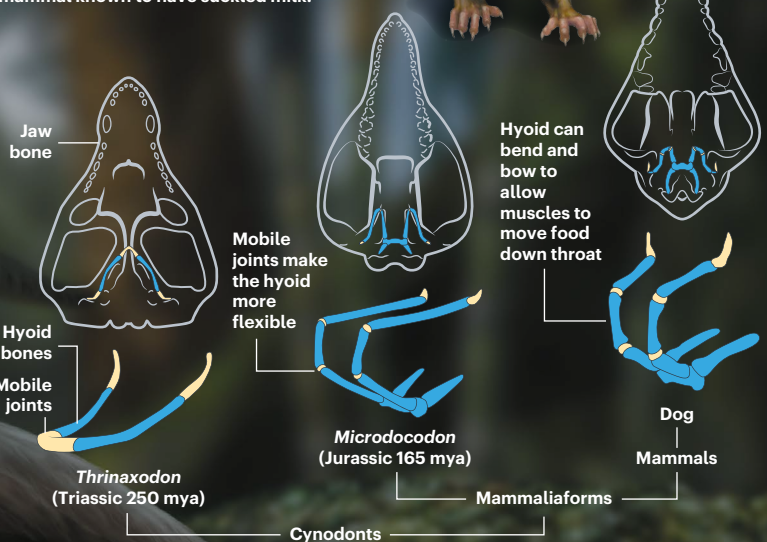
## SKY GLIDERS

Several modern mammals glide on wings of stretched skin, but the exquisitely preserved furry membranes of Jurassic-era *Maiopatagium furculiferum* revealed that this ability evolved early, by 160 million years ago. Squirrel-sized *Maiopatagium* probably feasted on fruit, but other gliders from the same period had teeth more suited to seeds.



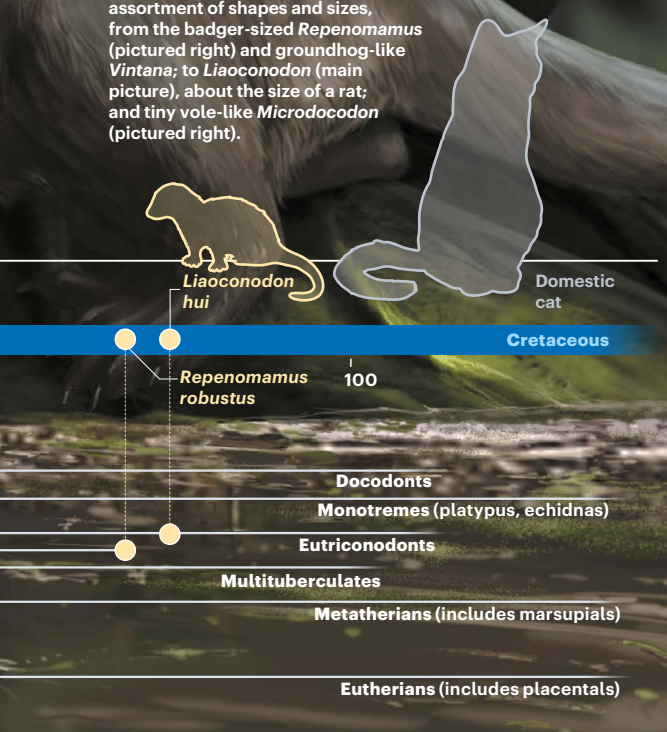

## SUCKLING AND SWALLOWING

All mammals nourish their young with milk, and to do this, the juveniles require a special bone in their throats to suckle and swallow. *Microdocodon gracilis* from the Jurassic has a form of this hyoid bone that resembles that of modern mammals and is the earliest mammal known to have suckled milk.




## GREAT AND SMALL

Early mammals came in a wide assortment of shapes and sizes, from the badger-sized *Repenomamus* (pictured right) and groundhog-like *Vintana*; to *Liaoconodon* (main picture), about the size of a rat; and tiny vole-like *Microdocodon* (pictured right).




## DINOSAUR DINNER

Species like *Repenomamus robustus* from the Early Cretaceous challenge the typical stereotype of early mammals as diminutive insect-eating fluffballs. This creature was the size of a badger and a specimen found in 2005 had the bones of a baby dinosaur – a beaked, bipedal *Psittacosaurus* – in its stomach.

ZHE-XI LUO/UNIV. CHICAGO

**This exquisitely preserved 160-million-year-old specimen of *Maiopatagium furculiferum* shows how early gliding evolved.**

▸ by cartilage. "The hearing function and the chewing function were still not completely separated," he explains. This was hard evidence of the evolutionary transition from jaw to ear.

Another unique trait of mammals is the sophisticated way they chew and ingest food in small parcels, rather than swallowing things whole as snakes and alligators do. To make that possible, mammals evolved a wide variety of complex teeth for biting and grinding food.

But as babies, mammals are nourished another way – by suckling from their mother's mammary glands. "Our whole group is named after this incredible biological innovation," says Luo. Drinking milk is made possible by the ability to suck and swallow, aided by the hyoid bones in the throat and muscles that support them. This apparatus also forms the voice box.

In July, Luo published a paper revealing a 165-million-year-old vole-sized docodont – a close relative of true mammals – that had the hyoid bones of its throat preserved[14]. *Microdocodon gracilis* is the earliest animal known to have been able to suckle like a modern mammal.

This level of detail is rare, and – similar to the study of the *Kayentatherium* hatchlings – the work on both the ear and throat bones has been made possible only through advances in microCT scanning techniques, says Krause. The technique has also revealed details about the olfactory abilities and brains of early mammals. These revelations are "breathing life into these early mammals in ways that were previously impossible and almost inconceivable", he says.

Much of the constellation of features we think of as defining mammals – complex teeth, excellent senses, lactation, small litter size – might actually have evolved before true mammals, and quite quickly. "More and more it looks like it all came out in a very short burst of evolutionary experimentation," Luo says. By the time mammal-like creatures were roaming around in the Mesozoic, he says, "the lineage has already acquired its modern look and modern biological adaptations".

## Family drama

Although the experts concur on many points, there is still much debate about how early mammal groups are related, and which groups are true mammals. That leads to uncertainty about how key traits evolved, says Hoffman.

One sticking point between Meng and Luo, who have each developed their own evolutionary trees, is the haramiyids. Meng thinks this early group belongs with true mammals, whereas Luo is convinced it's a side branch. The oldest known haramiyids are from 208 million years ago in the Triassic. If they are true mammals, then mammal origins date back at least that far – if not, then the oldest known mammal is 178 million years old, well into the Jurassic.

More fossils will help to resolve such questions, and bring more surprises. Krause and Meng say they are both studying exciting fossils, but are yet to publish their findings on them, and tens of unstudied specimens lie piled in the offices of their Chinese colleagues.

Palaeontologists have many items on their wish lists. One characteristic that Luo wants to understand is growth rates. Reptiles grow slowly throughout their lives, whereas mammals grow in bursts in youth and then plateau. He'd love to find a series of fossils from babies to adults to watch this happening. "That is one of the most critical features in mammals that help to define our biology," he says.

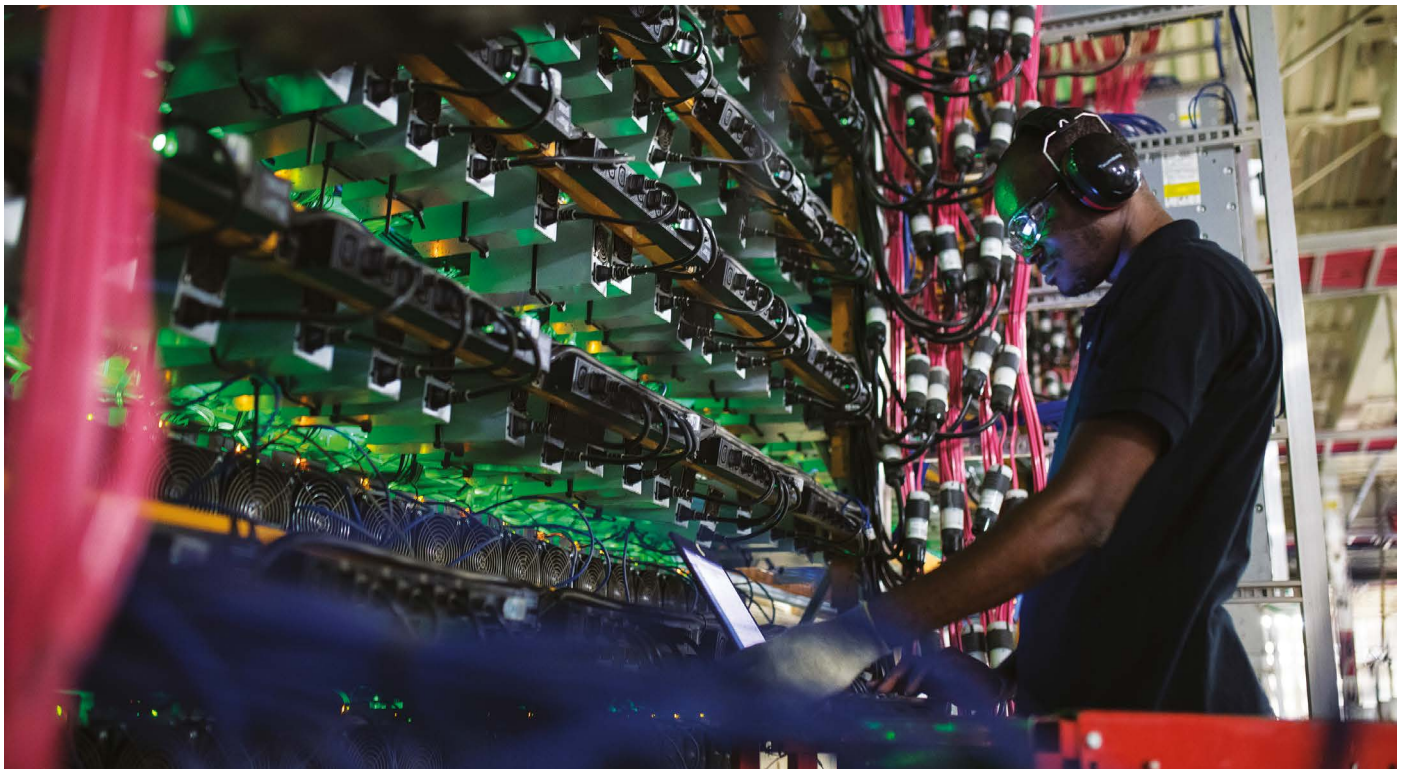
Both Hoffman and Meng agree that embryos and more babies would be significant finds – and, like the *Kayentatherium* discovery with its dozens of hatchlings, they would help us to pinpoint the date that mammal-style small litter sizes appeared. Meng's dream is to find a pregnant mammal. "This is always in my mind that I will find a mammal that inside the skeleton you can see some very delicate skeleton, which is either an egg that hasn't hatched, or it's a more interesting fetus," he says.

If the flurry of discoveries has taught researchers anything, it's that every fossil find has the potential to add a chapter to evolutionary history or even flip the prevailing narrative on its head. "We're really in this exciting, almost manic phase of lots of new evidence coming in, and it's going to take time to synthesize," says Brusatte.

**John Pickrell** is a science journalist and author in Sydney, Australia.


1. Hoffman, E. A. & Rowe, T. B. *Nature* **561**, 104–108 (2018).
2. Hu, Y., Wang, Y., Luo, Z. & Li, C. *Nature* **390,** 137–142 (1997).
3. Meng, J., Hu, Y., Wang, Y., Wang, X. & Li, C. *Nature* **444**, 889–893 (2006).
4. Luo, Z.-X. *et al. Nature* **548**, 326–329 (2017).
5. Meng, Q.-J. *et al. Nature* **548,** 291–296 (2017).
6. Meng, Q.-J. *et al. Science* **347**, 764–768 (2015).
7. Ji, Q., Luo, Z.-X., Yuan, C.-X. & Tabrum, A. R. *Science* **311**, 1123–1127 (2006).
8. Luo, Z.-X. *et al. Science* **347**, 760–764 (2015).
9. Krause, D. W. *et al. Nature* **515,** 512–517 (2014).
10. Hu, Y., Meng, J., Wang, J. & Li, C. *Nature* **433**, 149–152 (2005).
11. Luo, Z.-X., Yuan, C.-X., Meng, Q.-J. & Ji, Q. *Nature* **476,** 442–445 (2011).
12. Allin, E. F. & Hopson, J. A. in *The Evolutionary Biology of Hearing* (eds Webster, D. B. & Fay, R. R.) 587–614 (Springer, 1992).
13. Meng, J., Wang, Y. & Li, C. *Nature* **472**, 181–185 (2011).
14. Zhou, C.-F., Bhullar, B.-A. S., Neander, A. I., Martin, T. & Luo, Z.-X. *Science* **365**, 276–279 (2019).

# Books & arts



JAMES MACDONALD/BLOOMBERG VIA GETTY

**A technician monitors cryptocurrency-mining rigs at a Bitfarms facility in Saint-Hyacinthe, Canada.**

# The storied state of economics

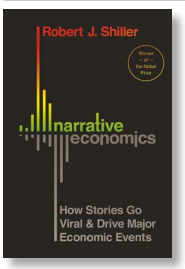
## Robert Shiller's study probes how social behaviour supersedes statistics in determining the fate of economies. **By Tim Jackson**

"Economists are tellers of stories and makers of poems," wrote the economic historian Deidre McCloskey in 1990. It's a curious observation for a profession that prides itself on hard-nosed, quantitative analysis and strives continually for predictive power. The Nobel-prizewinning economist Robert Shiller goes even further.

Stories are more powerful than statistics, he claims. The irrationality inherent in financial exuberance (and despair) defies the neat territory of numbers and demands a deeper excursion into the decidedly unruly world of narratives. That is the declared aim of his book *Narrative Economics*.

It's a compelling hypothesis. Since the 1960s, we have known that science is socially constructed. Since the 1980s, sociologists have sought to understand the 'social amplification of risk' – in which people are drawn inexorably towards stories of disaster or triumph (rather than statistics or probabilities) as the lodestone for the perceptions of risk that guide their everyday decisions. Around the same time, philanthropist George Soros adapted the concept of reflexivity to explain how investors' perceptions affect the social environment, which, in turn, informs their perceptions.

This feedback loop allows speculative bubbles to arise with alarming speed, and then collapse again. The phenomenon reached its apotheosis in a now infamous remark from Citibank chief executive Chuck Prince that "when the music stops, in terms of liquidity, things will be complicated. But as long as the music is playing, you've got to get up and dance." His prophetic words came just months before the 2007–08 financial crisis struck.




**Narrative Economics: How Stories Go Viral and Drive Major Economic Events**
Robert J. Shiller
Princeton University Press (2019)

Shiller elevates these insights into a full-blown exploration of the multiple ways in which narratives influence economic behaviour. Much as he tracked the rise and fall of asset prices in his Nobel-prizewinning work, he now charts the flux of narrative memes using Google's Ngram Viewer – which allows users to track the frequency of words and phrases in text over time – and Proquest's database of news citations. It's a quaint device, and there's a deceptive similarity between the time-series

graphics in *Narrative Economics* and those in his bestselling book *Irrational Exuberance* (2000). But the message is effective: the value of your story might go up as well as down.

The empirical core of the book is a detailed exploration of numerous real-life case studies, ranging from bimetallism (an old-fashioned form of money) to bitcoin (a brand-new one), and from the Great Depression of the 1930s to the Great Recession of recent years. Along the way, his anecdotes form a fascinating subscript. A convincing case is made, for instance, that fears of a 'singularity' – a point of no return arising from technological advances – are perennial. He notes numerous viral outbursts of this meme (associated with cotton mills, electricity and computers, for instance) dating back to the nineteenth century. Today's apocalyptic anxieties about a robot takeover are nothing new and should not be heeded, Shiller seems to imply. How that will turn out remains to be seen.

We learn that the mechanism through which a memorable turn of phrase goes viral can be described as a form of contagion, mirroring models from epidemiology. But we are also persuaded that viral success depends inherently on the messenger. Few remember that the phrase "the only thing we have to fear is fear itself", immortalized by US president Franklin D. Roosevelt during the Great Depression, was first uttered years earlier by economist Irving Fisher. It's troubling, of course, to be reminded that the rewards for creativity are often misallocated – particularly in today's plagiaristic world of social media, with its immense powers of narrative acceleration. But for me, this particular example raised a deeper concern.

Fear is a rational response from people whose livelihoods are under existential threat. So why would a president inveigh against it? The answer is that Roosevelt was painfully aware of the implications of fear. He was addressing what the economist John Maynard Keynes (borrowing from another long-forgotten creative) called the "paradox of thrift": the tendency of ordinary people to curtail their consumption in the face of economic uncertainty, and put their money into savings instead.

Such behaviour is sensible, admirable even, at the individual level. Perhaps it is so at the planetary level, too: lower consumption might benefit the environment. But economics has a problem with it. As people spend less, demand is suppressed, prolonging the recession. The same thing happened after the 2007–08 crisis. The paradox of thrift was the foundation for Keynes's most famous proposal: that governments provide stimulus that could return the economy to growth when people would not. This was the rationale for Roosevelt's New Deal package of reforms, and the inspiration for the proposed US legislation called the Green New Deal.

Keynes was a pragmatist; his prescriptions were a response to the diseases of the day. But he was also a visionary. In his essay 'Economic Possibilities for Our Grandchildren' (1930), he foresaw a time when our society would move beyond growth. It hasn't happened yet – in spite of economist Kenneth Boulding's remark to the US Congress in 1973 that "anyone who believes exponential growth can go on forever in a finite world is either a madman or an economist".

Shiller is clearly not a madman. But in the course of an otherwise fascinating exploration of the power of story, he never once acknowledges that eternal growth is itself just a narrative. He notes that the logic of relentless expansion conflicts with the logic of human anxiety. But he assumes that it is people who are at fault. Narratives can have clear, moral and prudential foundations, it seems, but they might still be cast as irrational.

**"It's troubling, of course, to be reminded that the rewards for creativity are often mislocated."**

Indeed, for Shiller, that memorable speech on the "fear of fear" shows government attempting to "lean against false or misleading narratives and establish a moral authority against them". Roosevelt's remark was designed to "create and disseminate counternarratives that establish more rational and more public-spirited economic behavior". What Shiller seems to be saying is this: when ordinary human sentiment runs counter to the prevailing logic of capitalism, the state must override it. It is a deeply suspect, potentially dangerous conclusion. But it, too, demonstrates just how pervasive narrative is.

Ultimately, *Narrative Economics* is an eloquent and accessible exposition of a seductive idea. It's a particularly compelling hypothesis for Britain, a country still reeling from a public referendum whose outcome was determined by viral confabulations of the most pernicious kind. We are all "tellers of stories and makers of poems". But neither economists nor politicians can claim moral authority over narrative truth. We must all choose carefully which stories we live by.

**Tim Jackson** is director of the Centre for the Understanding of Sustainable Prosperity at the University of Surrey in Guildford, UK, and author of *Prosperity without Growth*.
e-mail: t.jackson@surrey.ac.uk

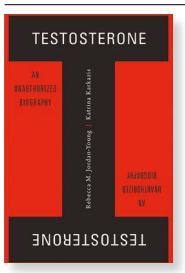
# Testosterone chronicles: truths and tall tales

## A book on the hormone dissects fact from fake and questions interpretations. **By Randi Hutter Epstein**

On 1 June 1889, renowned neurologist Charles-Édouard Brown-Séquard shocked his colleagues. Speaking at the Paris Society of Biology, the 72-year-old announced that a slurry made from the ground testicles of guinea pigs and dogs (injected under his skin ten times in three weeks) made him stronger. He also noted that his "jet of urine" lengthened by 25%.

Brown-Séquard was ridiculed by his peers throughout Europe for disseminating results with no scientific basis and promoting quack youth-enhancing 'cures'. Yet the bizarre elixir found favour with members of the public in the United States, United Kingdom and Europe – at least among men eager to recapture youthful sexual prowess. As the engaging book *Testosterone* explains, Brown-Séquard's testimonial helped to shape future studies that linked the hormone to alleged 'manliness'.

Anthropologist Katrina Karkazis and sociomedical scientist Rebecca Jordan-Young did not write *Testosterone* to rehash familiar tales of wacky hormone experiments of yore, although this is one of a few that they include. Their contention is that many testosterone researchers – then and now, and intentionally or not – interpret data with blinkers on. When the facts do not fit the paradigm, the authors argue, findings are moulded into flawed dogma. Karkazis and Jordan-Young strive to comprehend how scientific practice




**Testosterone: An Unauthorized Biography**
Rebecca M. Jordan-Young, Katrina Karkazis.
Harvard University Press (2019)

SIDNEY MOULDS/SPL

**Polarized-light micrograph of crystals of testosterone.**

around testosterone unfolds, and explore how the results "circulate and morph in the world".

Today, the biochemistry of this steroid hormone is well known, from its daily fluctuations to its synthesis from cholesterol and occasional conversion to oestradiol, a form of oestrogen. Testosterone is known to restore sex drive and muscle tone among men with ailments that reduce levels of the hormone, such as pituitary tumours. During puberty, a surge of testosterone in young men typically leads to enlargement of the muscles, penis, testes and prostate gland, and the emergence of secondary sex characteristics. In women, testosterone excreted by the adrenal glands and ovaries is generally important for ovarian function and bone strength.

Like pathologists doing a post mortem, Karkazis and Jordan-Young dissect the remains of a selection of studies. They parse statistics and the cultural context that prompted the research and influenced how the data were analysed. (Full disclosure: I have served on a history of medicine panel with Karkazis and, as medical authors writing about endocrinology, our paths have crossed several times.)

The authors delve first into testosterone's role in ovulation. The hormone and its precursor, DHEA, have a role in the maturation of ovarian cells; DHEA might boost fertility directly or as a mediator of oestrogen production. There are chapters focusing on traits often assumed to be associated with testosterone, such as athleticism. The authors also scrutinize the brouhaha surrounding a small psychology study[1] claiming that holding particular poses boosts testosterone production. There is a section on parenting, thanks to studies that created a fleeting media buzz by claiming that new fathers' testosterone plummets when they change nappies and do other nurturing chores[2,3]. And the authors discuss athletes who take testosterone to boost their abilities.

They do not dispute that injections, gels or patches that send testosterone levels skyrocketing above the norm build muscles when coupled with intense training. But they are sceptical about whether the hormone makes a large difference for every athlete. Some studies, they write, have found a correlation between high natural testosterone levels and speed and power; others show tenuous or no links. And a few studies link higher testosterone levels to worse performance.

Jordan-Young and Karkazis challenge murky definitions. They show how researchers define risk-taking through "weirdly narrow and also wildly divergent" behaviours, such as riding a motorcycle without a helmet. They cite a team that surveyed business students about their entrepreneurial experience and used a saliva sample to gauge their testosterone levels[4]. On the basis of these dubious data, the investigators concluded that those who had the highest levels, coupled with family business experience, were the most entrepreneurial.

When it comes to testosterone and aggression, the authors say that some of the most rigorous studies (double-blind, placebo-controlled) show no connection. What's more, they write that even the investigators of studies that tie testosterone to violence acknowledge that the link is inconsistent and weak. Yet the idea that testosterone drives violence remains widely accepted and "grossly overblown".

By setting the record straight, the authors build on their past record. Jordan-Young explored the evidence for putative neurological sex differences in the 2011 book *Brain Storm*; Karkazis demolished preconceptions about people who are intersex in her 2008 work *Fixing Sex*, which also explores the often disturbing history of 'treatments' for 'ambiguous' genitalia.

Although often academic in tone, the book is leavened by a welcome informality. The authors describe the link between testosterone and violence as a zombie: "a fact that seemingly can't be killed with new research". They personify testosterone as "T" and characterize their book as an "unauthorized biography". An authorized biography, they note, "sweeps away all kinds of details and smooths over contradictions". Theirs intends to pull back a veil that has obscured the field.

Still, I sometimes wanted more. In a chapter on ovulation, they quote a woman receiving fertility treatment who thinks that a therapy containing DHEA helped her to produce more eggs of higher quality. The authors note that the idea of testosterone aiding a woman's fertility has been anathema to reproductive endocrinologists, but quote only one clinician. That left me wondering whether other clinicians were still reluctant, or if this were part of standard treatment. I wanted to hear from other fertility clinicians.

In the opening of the chapter on athleticism, the authors refer to a 2012 meeting with an endocrinologist who explains that testosterone rises sharply in response to intense exercise, but that responses vary among athletes. Then, they describe an interview with a second expert who tells them the opposite, and also says that some types of sports training might lower testosterone. I wanted to know who these experts were.

Moreover, although Jordan-Young and Karkazis are lively storytellers, every now and then an anecdote doesn't jibe with the chapter's content. For example, they start the discussion on risk-taking with a delightful account of 63-year-old Annie Edson Taylor, who in 1901 went over Niagara Falls on the US–Canadian border in a pickle barrel. That seems a literary stretch: we know nothing of Taylor's hormonal state (except that because she was probably postmenopausal, her testosterone would have been low, and her oestrogen and progesterone certainly lower than before).

These quibbles, however, are minor in a deeply researched and thoughtful book that adds a fresh perspective to a growing body of work aiming to debunk myths about hormones.

**Randi Hutter Epstein** is writer in residence at Yale School of Medicine in New Haven, Connecticut, and author of *Aroused: The History of Hormones and How They Control Just About Everything* and *Get Me Out: A History of Childbirth from the Garden of Eden to the Sperm Bank*.
e-mail: rh152@columbia.edu

1. Carney, D. R., Cuddy, A. J. C. & Yap, A. J. *Psychol. Sci.* **21**, 1363–1368 (2010).
2. Gettler, L. T., McDade, T. W., Feranil, A. B. & Kuzawa, C. W. *Proc. Natl Acad. Sci. USA* **108**, 16194–16199 (2011).
3. Storey, A. E. *et al. Evol. Hum. Behav.* **21**, 79–95 (2000).
4. White, R. E., Thornhill, S. & Hampson, E. *Organ. Behav. Hum. Decision Process.* **100**, 21–34 (2006).

# Science by design: *Nature* renewed

## From custom typeface to digital-friendly logo, follow the journey to the journal's new look. **By Kelly Krause**

Should science be ugly? This is a serious question asked by serious people at seminars. Some assume that an aesthetically appealing presentation signals at best a lack of priorities, and at worst a lack of rigour.

I disagree. Science sorely needs best practices in visual communication as well as in information design, a mature field with quantitative methods. In my view, the idea that scholarly publishing should be divorced from evidence-based applications of good visual design is perplexing.

**"The 'flavour' of the typeface — the feelings it evokes, its personality — evolved over several months."**

Looking back over the past 150 years of *Nature*, we see an aesthetic that bends with time and trends, from ornate Victorian embellishments in 1869 to stark minimalism in the late 1960s. But design is not solely about how something looks; it is also concerned with how it works, and that understanding has never been more urgent than in the digital age. Design as a discipline exists to solve problems, and working researchers, readers and contributors have many. As publishers, we've asked how we might assist working scientists. We have heard your pleas, many stemming from information overload and the need to pack ever more data on to small screens.

So we are refreshing *Nature*'s look, and not just in honour of our 150th anniversary. We are in the early stages of our evolution towards designing for readers' digital reality. Here is a tour of what is different, and why and how we have changed it.

### Typography

A custom typeface, Harding, has been created for *Nature*'s new logo and much else: you're reading it right now. Harding is named after the late neurologist Anita Harding. Brilliant and generous, she published in *Nature* before she died in 1995 at age 42. According to colleagues, she was known for taking questions from the clinic back into the laboratory, and for her wry sense of humour. When she learnt that she had a terminal illness, she apparently joked that at least she wouldn't have to buy Windows 95.

Our team designed the typeface specifically for science, in partnership with Commercial Type founders Christian Schwartz in New York and Paul Barnes in London, and with London designer Mark Porter, whom *Nature* engaged for the overall redesign. Care was taken to identify the needs of technical material, because scholarly articles use classic type styles in unique ways. For instance, papers often have mathematical equations and formulae in the sub- and superscript lines, along with Greek letters and special characters. So we have made the sub- and superscript characters larger than standard, and created a Greek alphabet carefully honed to convey scientific meaning rather than typical Greek-language prose — for example, clearly rendering an alpha (α) in a shape that looks like a mathematical symbol, so that it is not easily confused with a Latin italic letter *a*. We have also made the italics more slanted so they are more distinct; single italic characters, such as *h* for Planck's constant, are often used as isolated symbols with scientific meaning.

Harding is designed to cope across the disciplines. It boasts an unusually large range of special characters, from triple prime and nabla to a full set of astronomical symbols and the 'click' phonemes found in some African languages.

A key consideration in Harding's overall design is performance on small digital screens. To boost readability in a limited space, it helps to enlarge the main portion of the lower-case letters, while making the ascenders and descenders (as in 'h' and 'g', respectively) smaller. Ultimately, this renders long, complex strings of words easier to parse, and allows for neat stacking of lengthy technical research-article titles over a number of lines.

The 'flavour' of the typeface — the feelings it evokes, its personality — evolved over several months. We initially looked at six fledgling concepts, each with distinct letterforms such as rounded serifs (the small strokes at the end of letters). After we winnowed these down to two, Harding emerged as the clear

## EVOLUTION OF *NATURE*

**The old logo**
A selection of logos from *Nature*'s past. The lower-case 'n' was introduced in 1974.


**A matter of contrast**
In typography, contrast refers to weight contrast, which is the difference between thick and thin strokes, shown lower right. *Nature*'s outgoing logo is very high contrast, which is not suitable for digital, particularly mobile, devices.


Areas of high contrast (circled) tend to blur and disappear in digital environments.

**Weight contrast**


**Designing a typeface**
Harding's unusually high x-height is engineered for optimum legibility at small sizes, particularly on mobile screens, but also in print. The short ascenders and descenders help to stack long strings of words into small spaces.


**Greeks** – designed for scientific use
$\alpha\ \lambda\ \mu\ \nu\ \pi\ \tau\ \varphi\ \psi$

**Sub- and superscript** – larger than most fonts
$e^{i\pi}+1=0$

**Italic** – more slanted for mathematic or scientific terms
$p^2+2pq+q^2=1$

**Astronomical symbols**
☉ ☿ ♀ ♁ ♂ ♃ ♃ ♅ ♆ ♇

**Evolution**
*Nature*'s updated logo is based on our new typeface, Harding. In this illustration, the standard Harding letters are light blue, with the new logo type in dotted black outline.


For the logo, we added a ball terminal to the 'r' and a tail to the 'a', in a nod to logos of the recent past.

**Final logo**


winner. We aimed for an overall impression of calm, rational intelligence with perhaps a dash of British formality and wit.

The myriad design considerations behind *Nature*'s new typeface serve one goal: to improve the reading experience for researchers and policymakers globally, and enhance comprehension and insight.

### Logo

*Nature* has had at least ten logos since 1869, reflecting the styles of successive eras (see 'The evolution of *Nature*'). All, up to now, were designed for print. The relatively large, luxurious physical space of a printed cover allows for fine detail in a way mobile devices do not. Finely worked features, digitized into pixels, can look fuzzy on smartphone screens.

Using the Harding typeface as a basis, we have updated the logo. Weight contrast – the variation of thin and thick lines in a letterform – was an important factor because high contrast aids pixellation on the small screen. For the logo, we modified Harding slightly, adding a tail to the letter 'a' and a rounder terminal on the 'r', to align it with recent versions of the logo. There's an echo of the Baskerville Old Style logo from the early 1970s, but engineered for digital performance. And our team has retained the democratic lower-case 'n' that has been in use for almost 50 years.

We have also prioritized digital platforms by simplifying all *Nature*-branded journal logos. This was a particular challenge for journals with very long names, such as *Nature Structural and Molecular Biology*. Because social-media channels have grown in importance, we have also created a system of abbreviated forms for the tiny avatars on those platforms.

### Colour

Perhaps the most radical change to *Nature*'s look is the removal of the red bar from the top of its web page. The journal design has incorporated red only since the late 1990s. (Before that, orange persisted in the logo and printed pages for four decades.) *Nature*'s 'red period' was intertwined with the rise of the web, but as digital design language has matured, red is now often associated with unpleasant online features such as error messages. More importantly, by removing the red, we help content to stand out more cleanly. Research shows that elimination of unnecessary colour elements eases cognitive load.

All these elements – typography, logos and colour – form the basis of *Nature*'s new design language across digital, print and anywhere you might find us, from coffee table to Twitter feed. This language will most certainly evolve, driven by researchers' needs.

**Kelly Krause** is *Nature*'s creative director.

Engineer Adriano Olivetti in his typing machine factory in Ivrea, Italy.

KEYSTONE-FRANCE/GAMMA-KEYSTONE VIA GETTY

# Turbulent birth of the personal computer

The strange circumstances surrounding the invention of the world's first PC are probed by a new book.

**By Sharon Weinberger**

In the depths of the cold war, an Italian industrialist on the cusp of marketing the first personal computer dies on a train to Switzerland. Adriano Olivetti has had contact with Western spy agencies; his associates hint that his heart attack might not be what it seems. Such is the thriller-esque start to biographer Meryle Secrest's *The Mysterious Affair at Olivetti*.

At the heart of Secrest's book lie two questions: how did the Italian typewriter company Olivetti produce the world's first PC in the 1960s — long before its competitors — only to have its work fall into obscurity? And could Adriano Olivetti's death be linked to the company's disappearance from computer history? Secrest weaves a startling narrative around these events, involving a US intelligence agency and an information-technology multinational.

The story goes back to Camillo Olivetti, the Jewish-Italian industrialist who founded the company in Ivrea, Piedmont, in 1908. His visionary son Adriano, who succeeded him as company head in 1938, was interested in architecture, politics and technology. He began to look beyond typewriters to machines combining the best aspects of form and function. More crucially, he started to expand from mechanical typewriters into electronics.

When the Second World War broke out, Adriano Olivetti paid lip service to the fascists while secretly working to remove prime minister Benito Mussolini, all while keeping his factory going and his family alive. He survived the war, the company thrived, and he opened an electronics laboratory that drew on his experience in the United States. In the late 1950s, the company created one of the world's first transistorized mainframes, the ELEA 9003.

Olivetti's death in 1960 threatened to derail the plans he had set out for the company to further expand into computers. Moreover, the firm was in a downward spiral, following his decision in 1959 to buy his main competitor, the US typewriter firm Underwood. Yet Mario Tchou, a key engineer who oversaw the company's electronics work, was already thinking about shrinking mainframes into something that could sit on a desk. Adriano's talented but less savvy son Roberto oversaw manufacturing of the Programma 101 (P101) desktop computer, which made its debut in 1965. It was the world's first PC, and sold an astonishing 44,000 units over several years, including some to NASA. But the company's computer manufacturing was eventually overtaken by its competitors, particularly in the United States.

That sounds like the guts of a great technology history. The book's subtitle, meanwhile, promises spy-versus-spy intrigue involving the CIA and US computer giant IBM. However, Secrest focuses more on the Olivetti family than its products. There is a bare-bones description of the P101 and how it was developed: the

ULLSTEIN BILD/GETTY

**Olivetti's table-top computer, 'Programma 101' in 1966.**

programming system, Secrest notes, "took an enormous amount of experimentation". But there is little more on what must be an intriguing techie history.

Secrest thus also misses several opportunities to tease out intriguing storylines. For example, Adriano Olivetti's insistence that something sitting in your office should be both functional and beautiful almost certainly inspired Apple co-founder Steve Jobs. The aesthetic similarities between Olivetti's 1960s-era showroom on Fifth Avenue in New York City and today's iconic Apple stores are uncanny.

The book's treatment of espionage is at times more detailed than its take on tech. Secrest describes fascinating wartime contacts between Adriano Olivetti and British and US intelligence agencies. While feigning loyalty to the Fascist Party, the industrialist was secretly meeting with the US Office of Strategic Services (OSS), the predecessor to the CIA, which dubbed him 'Agent 660'. There is drama in this. But as Secrest makes clear, Adriano was no 007; the OSS never acted on his plans, and British intelligence seemed to dismiss him as a dreamer and deemed his convoluted scheme for toppling Mussolini unrealistic.

The narrative takes a stranger turn around Adriano Olivetti's death. It seems plausible that, saddled with mounting debt and Underwood's outdated factories, a 58-year-old businessman might die of a heart attack. Instead, Secrest decides that the CIA murdered Olivetti — as well as Tchou, who died in a car accident in 1961.

Gaining access to CIA records is certainly arduous, and Secrest describes her unsuccessful attempt to meet with the agency's historian, David Robarge. In the absence of insider insights or access to fresh archival records, she turns to a car-repair shop owner in Rockville, Maryland, for confirmation of her theory that the CIA engineered the car accident that killed Tchou.

The CIA, of course, really has attempted to assassinate certain figures, such as Patrice Lumumba, the first prime minister of the Democratic Republic of the Congo. But Secrest presents no evidence that US spies were involved in Olivetti's death. She implies that IBM, too, was somehow implicated, citing cold-war competition and the company's work for the US government and intelligence. (She reminds us that IBM, as documented by Edwin Black in his 2001 book *IBM and the Holocaust*, sold technology to the Nazis in the 1930s.) A link between that and the Olivetti affair is never aired, however.

**"A more interesting historical question is why US computer science advanced so quickly during the cold war."**

This conspiracy-mongering is a shame. Secrest does all the right research, and the clues to the company's troubles (and Olivetti's woes) are right in front of her. In an era of rampant conspiracy theories, such as bizarre allegations involving the Jewish Hungarian-American billionaire George Soros, we rely on scholarship to pull out the facts, not just the speculation.

A more interesting historical question is why US computer science advanced so quickly during the cold war, leaving Europe behind for decades. It's likely that this happened because the Pentagon and US intelligence agencies invested in companies and technologies that had no immediate commercial prospects, but served US strategic interests (see page 481). The relationship between spies, soldiers and computer scientists during and after the second half of the twentieth century is worthy of serious exploration. *The Mysterious Affair at Olivetti* does not offer that.

Yet this book is, in other ways, a laudable attempt. It shines when describing Adriano Olivetti's interest in architecture (Secrest authored the 1992 book *Frank Lloyd Wright: A Biography*). Secrest writes well on the aesthetics of Olivetti machines and Adriano's attraction "to clean, boxy lines", the signature of the Bauhaus movement. Her biographer's instinct — choosing a visionary figure whose contributions have not been fully appreciated — is also to be applauded. As she laments, "the Programma 101 has not been well served by computer historians on or off the Internet." She is right. That record remains to be filled.


**The Mysterious Affair at Olivetti: IBM, the CIA, and the Cold War Conspiracy to Shut Down Production of the World's First Desktop Computer**
Meryle Secrest, Knopf (2019)

**Sharon Weinberger** is the author of *The Imagineers of War: The Untold Story of DARPA, the Pentagon Agency That Changed the World*.
e-mail: sharonweinberger@gmail.com

# Comment

ILLUSTRATION BY SEÑOR SALME

# Can marketplace science be trusted?

**Paul Lucier**

Historian Paul Lucier traces the explosion and fragmentation of industrial research in the fifth essay in a series on how the past 150 years have shaped today's science system.

**Anniversary collection:** go.nature.com/nature150

Four years after the first issue of *Nature* was published, the US National Academy of Sciences (NAS) faced an existential crisis. In October 1873, one of its original members demanded the expulsion of another member for swindling. Josiah Whitney, California's state geologist, accused Benjamin Silliman Jr, professor of applied chemistry at Yale University in New Haven, Connecticut, of accepting large sums from California oil companies in return for favourable, possibly fraudulent, science. Silliman responded forcefully that company funding for science was evidence of responsibility, not misconduct: companies needed objective "technical opinions". Without science, swindling would be more common, he argued.

NAS president Joseph Henry, secretary of the Smithsonian Institution and a former consultant to Samuel F. B. Morse, inventor of the telegraph, had to agree. If the NAS expelled every member who had ever consulted for a private company, it would not survive. Henry rejected the efforts to remove Silliman. More importantly, he resolved to expand the NAS membership; new members were to be judged on the basis of their research, not on the source of their income[1]. By the 1870s, it was already clear that industry relied on science.

The Silliman–Whitney controversy marked a watershed in the relationship between science and industry. For US scientists, as well as many in Britain and Europe, private companies had become valuable patrons, supplying both funds for research and problems to be researched, and were gainful employers who provided short-term commissions. Likewise, companies regarded scientists and their findings as profitable to the development of their respective industries.

Over the next 150 years, relations between science and industry continued to evolve — in four significant stages. Scientists moved from part-time consultants to full-time corporate researchers, and then to academic entrepreneurs. Industry grew from a scattering of local businesses to a concentration of large companies, and on to multinational corporations with global reach. Although these transformations might seem symbiotic, and even inevitable, the very fact that US scientists and industries emerged as leaders and exemplars (in terms of employment, funding, publishing, patenting and innovating) serves as a cautionary reminder of the contingent nature of such developments.

## Consultancy (1820–80)

At the heart of the NAS crisis was an essential tension in the relations between science and industry: can the pursuit of knowledge be corrupted by the pursuit of profit? To Whitney and his allies, the answer was obviously yes. Their 'pure' science needed to be practised in places protected from the profit motive, such as government agencies or well-endowed universities. Silliman and supporters of 'applied' science, by contrast, believed the interactions between science and industry to be mutually advantageous. Indeed, the emergence of a distinct kind of endeavour called applied science characterized a new era in which research would address more and more industrial concerns, and private enterprise would, ideally, become a steady supporter of that work[2].

The profession of scientific consulting goes back to the early nineteenth century, when individuals or groups of capitalists occasionally commissioned scientists to examine prospects in farming, mining, transportation (canals and railroads) and manufacturing. These fee-for-expertise engagements were short term and advisory. By the 1870s, changes in US commercial law (similar to those in British and European law) allowed the formation of limited-liability, joint-stock companies. These businesses, with their large pools of funds and numerous shareholders looking for investment assurances, regularly consulted scientists. As the engagements became both more routine (continuous testing and analysing of existing products and processes) and more investigative, scientists began to receive lucrative contracts and retainers[1].

In the United States, geologists were among the most active consultants during the Gilded Age, a period of rapid economic growth from the 1870s to the 1890s, especially in precious-metal mining in the area west of the Mississippi River. In Britain and Germany, the most prolific consultants were chemists, because of their essential expertise in new products such as acids, soaps, paints and especially synthetic dyes, including mauve and alizarin. Consulting chemists also found themselves in prominent public roles as expert witnesses in sensational patent cases. Witness-box quarrelling among chemists made good newspaper copy, and it highlighted profound developments in the chemical industries. Changes in patent law in the United States, Britain and Germany allowed inventors to claim those new chemical products and processes as their intellectual property (IP) instead of judging them to be scientific discoveries, which were, by definition, unpatentable.

## Industry (1880–1940)

At the turn of the twentieth century, the independent consulting scientist was replaced by the salaried researcher in new industrial laboratories. These labs represented the incorporation of applied science; that is, the creation of a separate place within the organization for 'research and development' — a phrase that entered the lexicon at this time.

In Germany, the largest dye companies, such as Bayer, Hoechst and BASF, were the first to establish dedicated labs for chemical research. These were connected to production departments, also staffed by university-trained chemists, and to specialized legal departments, from which the new products and processes were submitted for patenting. This type of industrialized invention, with close connections between German academic chemistry and company labs, was firmly established before the First World War[3].

In the United States, the prototype for the industrial research lab appeared in the electrical industry, when inventor Thomas Edison set up an 'invention factory' in Menlo Park, New Jersey, in 1876. Edison wanted to replace what had been an unpredictable act of creative genius with a regular and reliable system. He recruited machinists, mechanics, chemists, physicists and mathematicians to work on technical problems connected to telegraphy and electric lighting. Although their efforts were collaborative, only the 'Wizard of Menlo Park' (the singular inventor) was listed on more than 1,000 US patents, including those for the phonograph (1878) and electric light bulb (1880)[4].

**"Having research in thrall to industry raised the alarm, again, that capitalism corrupted science."**

The looming expiration of that original light-bulb patent and the threat from other lighting companies impelled General Electric (GE), the corporation that took over Edison's Electric Light Company and all his patents, to establish the aptly named Research Laboratory in 1900 in Schenectady, New York. This proved profitable within a decade — commercially, with the invention of a new light bulb that restored GE to its dominant market position, and professionally, with the recruitment of more than 250 engineers and scientists.

A few other large US corporations followed suit and pioneered their own formal research and development (R&D) labs — DuPont (1903), Westinghouse Electric (1904), American Telephone and Telegraph (AT&T, 1909) and Eastman Kodak (1912).

It was the First World War and the embargo on all German products, especially chemicals, that was the catalyst to the golden age of 'industrial research', a neologism of the 1920s. Between 1919 and 1936, US corporations established more than 1,100 labs in nearly all industries — petroleum, pharmaceuticals, cars, steel — thereby dominating the world's industrial research. In 1921, these employed roughly 3,000 engineers and scientists; by 1940, there were more than 27,000 researchers. At the end of the Second World War, the figure was nearly 46,000 (ref. 5).

This remarkable proliferation reflected the massive scale of vertically integrated corporations that controlled nearly all areas of their respective industries, from natural resources through R&D to mass production and mass marketing. Industrial research was also fuelled by radical changes in US patent law that allowed these behemoths to claim the IP of their employees. The inventor was now the corporation.

During the Great Depression, critics singled out modern big business for its ruinous consequences to society — unemployment, overproduction and bankruptcy. Having research in thrall to industry raised the alarm, again, that capitalism corrupted science. So corporate captains and R&D directors marshalled the cornucopia of wondrous consumer products ('technology' in the new parlance) created by their science-based industries. In this story, science in industry was good; it guaranteed efficacy, efficiency and safety. In words that nineteenth-century consulting scientists would have understood, consumers could trust these modern technologies (and their corporations) because of the R&D.

At the World's Fair in New York City in 1939, industry paraded the fruits of its science. The Radio Corporation of America (RCA) introduced consumers to the television. International Business Machines (IBM) showed off its electric typewriter. GE exhibited its new electrical refrigeration system, and DuPont, under its banner "Better Things for Better Living through Chemistry", showcased a synthetic fibre called nylon[6].

Fears of corporate corruption of science were put to rest by awards of the Nobel prize. In 1931, two Germans, Carl Bosch and Friedrich Bergius, became the first industrial researchers to win in chemistry. The next year, GE's Irving Langmuir won the chemistry prize, and in 1937, Clinton J. Davisson of Bell Telephone Laboratories (Bell


BETTMANN

**US firms paraded the fruits of their industrial research at the 1939 World's Fair in New York City.**

Labs) won a share of the Nobel Prize in Physics.

The largest research facility in the United States was Bell Labs, established in 1925 in New York City to consolidate the R&D arm of AT&T and Western Electric, its telephone-manufacturing arm. The labs had around 3,600 staff members and a budget in excess of US$12 million. (GE allocated less than $2 million to its Research Laboratory.) The first president of Bell Labs was the physicist Frank Jewett. In 1939, he became the first industrial scientist to be president of the NAS[7].

In short, national standing and international acclaim seemed to confirm that science done under the auspices of industry was equal to science in universities or governments. Still, industrial labs of the 1920s and 1930s were not simply universities without students. As institutions of applied science, they always needed to show corporate headquarters their value in terms of profitable products and processes.

## Military (1940–80)

By the time the New York World's Fair closed in October 1940, Europe was already at war. The United States entered in December 1941, and the Second World War transformed the relationship between science and industry, along with the very terms — and even the history — of those relations.

The prime mover in all those changes was the US military and the unprecedented amounts of money it allocated — through new forms of contracting and subcontracting — to scientific research. During the war, the Office of Scientific Research and Development, under its director Vannevar Bush, signed more than 2,300 research contracts, worth roughly $350 million, with more than 140 academic institutions and 320 companies. About two-thirds of that funding went to universities; the Massachusetts Institute of Technology (MIT) in Cambridge, for example, received more than $200 million for its Radiation Laboratory for research on radar. Corporate R&D also received unrivalled amounts: AT&T was allocated $16 million, GE $8 million and RCA, DuPont and Westinghouse between $5 million and $6 million each[8].

But by far the most prodigious investments in R&D flowed from the War Department ($800 million) and the Navy Department ($400 million). The largest portion of that went to private industry ($800 million), much of it directed towards emergent industries with compelling national-security interests — for example, aerospace, electronics, computing and nuclear technology[8].

The US military had not intended to become the commander-in-chief of US science, but by the end of the war it was apparent, at least to Bush, that the federal government needed a plan. In his 1945 report to US president Franklin D. Roosevelt, *Science — The Endless Frontier*, Bush presented a vision for US science policy that would guide and define both university science and corporate R&D throughout the cold war. The endless frontier was 'basic' research, the kind performed "without thought of practical ends", a direct throwback to the nineteenth-century idea of pure science. The US military would fund this to boost industrial research because, the reasoning went, basic research was "the pacemaker of technological progress".

Here, then, was a new argument. As many commentators at the time and since have pointed out, it did not reflect either the experience of the war years (during which multifunctional teams worked on military projects such as the atomic bomb or radar) or of the previous decades (in which multifunctional teams worked in R&D labs on corporate projects such as the light bulb). *Science — The Endless Frontier* thus propounded a different idea for developing new technologies, both military and commercial. In time, this became known as the linear model of innovation[9].

The theory posits a conveyor belt, beginning with basic science and moving smoothly along to development, then to manufacturing and production, and culminating with technology or innovation. Increase the amount of basic science and the (alleged) result would be more technology, innovation and overall economic growth. Theoretically, basic research was to be centred in universities (and military funding did transform US universities and their science departments accordingly). But corporate R&D labs were also contracting with the military, as they had been during the war. With these military contracts, as well as enlarged funding from corporate headquarters (business leaders also bought into the linear model), industrial labs were redirected away from applied science and towards basic research[10].

Such faith in endless scientific innovation combined with prodigious financial resources led to the creation of central corporate research labs. These functioned more or less independently, which nicely suited the new organizational structure of multinationals. In place of vertical integration, sprawling conglomerates adopted horizontal organizational structures comprising multiple divisions (the M-form organization), in which each division, including the central research lab, operated on its own.

Leading research labs relocated to the countryside, far removed from headquarters and any connection to manufacturing. RCA Laboratories Division, for example, expanded its campus near Princeton, New Jersey, after 1945 and started work on colour TV and semiconductors. In 1956, Westinghouse built up its research labs in Churchill outside Pittsburgh, Pennsylvania, for nuclear research. IBM set up its Thomas J. Watson Research Center, designed by the modernist architect Eero Saarinen, in Yorktown Heights near New York City in 1961, to work on lasers, semiconductors

and other computer-related physics. And Bell Labs moved its research headquarters to Murray Hill, New Jersey.

At its height (before 2001), Bell Labs conducted world-class research in many fields (physics, mathematics, radio astronomy) at numerous sites. Its largest campus at Naperville near Chicago, Illinois, employed 11,000 people. The 191-hectare flagship campus at Holmdel, New Jersey, some 30 kilometres south of New York City, included a magnificent mirrored-glass building also designed by Saarinen in 1962.

These 'industrial Versailles' did research without much development; they had indeed been converted into universities without students[11]. As industrial ivory towers, they hoovered up university faculty members and PhD scientists and engineers, promising them time and resources to pursue their own agendas, and offering them open publication policies that allowed their results to appear in the most prestigious journals. By the mid-1950s at RCA in Princeton, half of the staff were theoretical scientists and more than 75% of the contracts were with the military. DuPont, likewise, increased its scientific staff by 150% in the decade after the war, with the greatest growth in fundamental chemistry being at its Experimental Station near Wilmington, Delaware. By the early 1960s, the number of engineers and scientists employed in US industrial research topped 300,000 (ref. 12).

These leading corporate laboratories — Bell Labs, IBM, Westinghouse, DuPont, RCA (Princeton), Xerox Palo Alto Research Center (PARC, 1970) — became powerhouses of basic science. Between 1956 and 1987, 12 corporate scientists won Nobel prizes. Bell Labs alone has collected eight in physics and one in chemistry since the Second World War, including one for its most famous technology, the transistor, in 1956. In the early 1960s, corporate researchers authored 70% of papers appearing in *Physics Abstracts*. By 1980, Xerox PARC matched the world's leading universities on citation impact[6,8].

With its emphasis on basic science as the necessary prerequisite to any future technological progress, the linear model was a break with the past. It prompted a new interpretation of the historical relations of science and industry. In the 1950s and 1960s, economists, historians and other scholars began to re-examine the latter half of the nineteenth century, and claimed to have discovered a 'Second Industrial Revolution'. Characterized by the chemical and electrical industries, this revolution involved replacing the old trial-and-error methods of invention used in the dirty industries of the 'First Industrial Revolution' (textile factories, coal mines and iron foundries) with science-based methods. In this revisionist history, glamorous synthetic dyes and bright electric bulbs sprang directly from the pure science of organic chemistry and electromagnetic physics. History thus seemed to provide definitive evidence for the necessity of continued funding of basic science, as well as a ready explanation for why US and Western European corporations had dominated the world's economy for more than a century[13].

It was not to last.

## Outsourcing (1980 on)

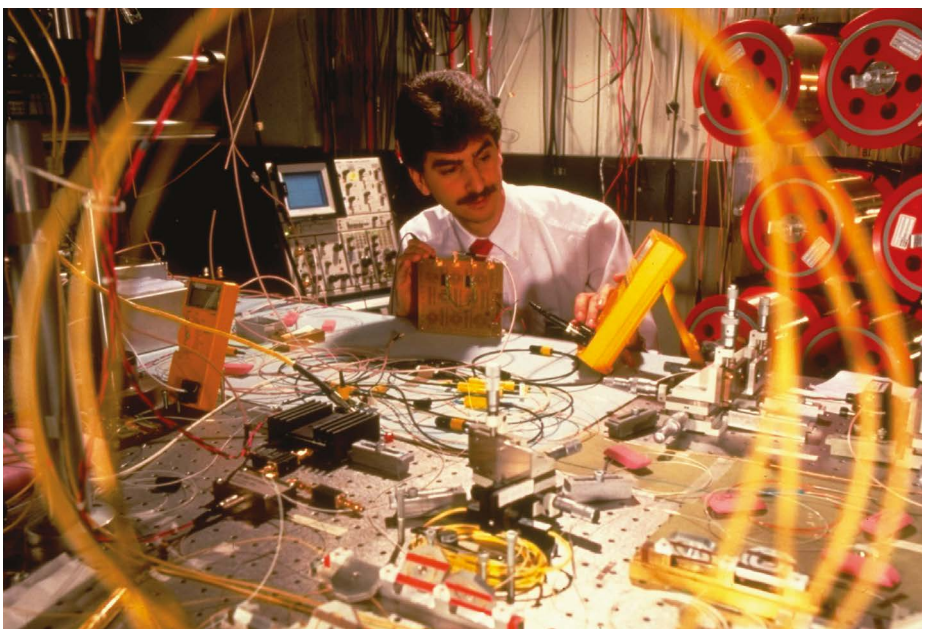
Corporate investment in basic science had been sustained by dominant positions in international markets. AT&T, DuPont, IBM, Kodak and Xerox held more than 80% market shares in their respective core businesses. Then the oil shocks of the 1970s, combined with widespread stagflation (high inflation, slow growth), weakened the US and European economies. Global competition increased, especially from Japanese and South Korean firms. In the early 1980s, growing free trade squeezed profit margins even further.

In response, US corporations began to restructure and downsize. Business leaders and shareholders decided that the multi-division conglomerate had become too unwieldy to compete. A new, leaner corporation was required. One way to restructure was outsourcing, replacing internal suppliers with external ones. Corporations began to relocate their manufacturing, once the backbone of the industrial economy, to plants in lower-cost and less-regulated countries. (The pace has only accelerated, especially after 2001, when China joined the World Trade Organization.)

Another way to downsize was divestiture, selling off subsidiaries unrelated to the core business. To shareholders seeking quick profits, long-term corporate research looked like a financial liability. The central laboratory became a prime target. In 1988, RCA sold off its Princeton lab as an independent business, Sarnoff Corporation. In 1993, IBM slashed $1 billion — roughly 20% — from its R&D budget. The German corporation Siemens bought Westinghouse's Churchill laboratory in 1997, and in 2002, PARC, the former division of Xerox, became an independent company. In 1996, AT&T, following the break-up of its phone monopoly, spun off the vaunted Bell Labs as a separate company, Lucent Technologies (in 2016 this was taken over by Nokia, the Finnish telecommunications company). The Holmdel campus closed in 2007. Within a year, just four scientists remained at Murray Hill doing fundamental physics research. It was the end of an era[14].

Accompanying globalized competitive markets, liberalized free trade and shareholder short-termism, the US military began to cut back funding for basic science at corporate labs. With the exception of a few years in the early 1980s (US president Ronald Reagan's Strategic Defense Initiative, the 'Star Wars' programme), the US government steadily reallocated research funds to universities and other non-profit organizations, particularly towards medical schools and research hospitals through the National Institutes of Health (NIH). With continuous funding, new fields (molecular biology, biochemistry and biotechnology, for instance) surged past the diminished physical sciences. By 1988, only about 10% of basic research articles in physics were authored by industrial scientists; by 2005, the number had plummeted to less than 3% (ref. 15).

The demise of the corporate research lab heralded the death of the linear-model idea. Many scholars concluded that it was too simplistic. The pathway from science to technology was neither straight nor singular, and perhaps not even one way (technological advances can



OVAK ARSLANIAN/THE LIFE IMAGES COLLECTION VIA GETTY

**Bell Labs in the 1990s: a researcher testing data transmission through fibre-optic cable.**

also lead to scientific discoveries). For corporate executives, investment in basic science did not seem to pay off. DuPont discovered no new nylons; Kodak failed to produce a revolution in photography; RCA lost its edge in consumer electronics; IBM ignored the personal computer; and Xerox PARC let slip the graphical user interface.

In the late 1960s and 1970s, small firms such as Intel, Microsoft, Apple, Sun Microsystems and Cisco Systems did commercialize the basic research being done at the larger corporations. Without establishing traditional research labs of their own, these players came to dominate the new information technology (IT) industry. In 1991, for example, when Microsoft created Microsoft Research — one of the largest industrial labs of its generation — its declared mission was not basic science, but innovation. In a more extreme case, Apple co-founder Steve Jobs shut down a fledgling research lab in 1998 in the belief that innovation would not require any investment in R&D.

Until 2010 and the emergence of machine learning, artificial intelligence (AI) and the Internet of Things, most technology companies ignored basic research. In 2012, following Jobs's death, Apple began investing in R&D again, particularly in AI. Likewise, Amazon, Google, Facebook and Uber began to recruit AI researchers from academia. This brain drain has become so serious that universities have begun to worry about their ability to train future AI researchers.

Twenty-first-century corporations value science (particularly, patentable discoveries) and still think that basic research can lead to invention and innovation. They would just prefer that someone else do it (and pay for it). In business terms, they optimize their 'supply-chain management', a phrase that gained currency in the 1990s, by replacing stable in-house labs (warehouses of scientists and engineers) with flexible contract research. Their ability to do so was greatly facilitated by the US government and the loosening of antitrust enforcement. The settlement of the monopoly case against Microsoft in 2001, for example, stands in stark contrast to the forced break-up of AT&T in 1984.

Moreover, the US government now permitted innovative start-ups to acquire new technologies, patents and licences from other companies and independent non-profit organizations such as Sarnoff and PARC, and to engage in extensive collaborative research with institutes and universities. Microsoft Research, for instance, now has labs around the globe (New York City, Beijing, Bangalore) and on several university campuses (MIT, the University of California, Santa Barbara, and Cambridge, UK), which account for 20% of patents in AI worldwide. Google, by contrast, mostly underwrites academic research through grants, fellowships, internships and visiting positions.

Universities have traditionally been the home of basic science. In the twenty-first century they have also become the source of innovation and entrepreneurship, in part because of sweeping changes in US patent law. In 1980, the US Supreme Court (in *Diamond* v. *Chakrabarty*) significantly expanded what could be patented to include new life forms. That same year, the US Congress passed the Bayh–Dole Act, permitting universities to patent the results of research funded by the NIH or other federal agencies and conducted on their campuses by faculty members, students and employees. Universities started filing for patents at an increasing rate — from 2,266 in 1996 to 5,990 in 2014. The university is now an inventor[16].

The most prominent industry that has been transformed by these legal and policy changes has been biotechnology. In 1976, a university biochemist and a venture capitalist founded Genentech, the first biotech firm. Genentech focused, as did other biotech start-ups (Amgen in 1980 and Genzyme in 1981), on translating basic science done in universities and, subsequently, in-house into patents and other forms of profitable IP. They facilitated that linear movement from research to development. Further commercialization towards the manufacture and distribution of drugs and therapies was taken up by traditional big pharmaceutical corporations. Eli Lilly (founded in 1876), for example, guided Genentech's first drug (synthetic human insulin) through clinical trials and brought it to market[17].

**"The pathway from science to technology was neither straight nor singular, and perhaps not even one way."**

The emergence of biotech represented both a new business plan (entrepreneurial scientists partnering with venture capitalists to sell their research) and a new model of innovation. Here, industry shifted from a single internal or closed source of research to multiple external or open sources[18]. In this model, academic entrepreneurs, commercialized universities, globalized contract-research institutes and numerous small research start-ups supply the science and the IP. Larger, more established firms then develop and commercialize these into new products and processes.

According to some economists and business scholars, open innovation characterizes a 'Third Industrial Revolution'[19]. From their perspective, the university professor seeking to patent the results of federally funded research to form a start-up, with seed money from venture capitalists, is the direct descendant of the consulting chemist of the nineteenth century. In this ecosystem, a population of nimble researchers and small firms has displaced a pack of lumbering corporate labs[20]. To critics and less-sanguine academics, the twenty-first-century relations of science and industry illustrate the commodification of university research and the corruption of the pursuit of knowledge by the profit motive[21].

Today, a complex innovation web has replaced the old conveyor belt. This is another new model — global commercialization. Supply-chain science is premised on the belief that research is a fungible commodity to be bought on demand and sold by the lowest-cost lab. In some ways, twenty-first-century contract research is reminiscent of nineteenth-century consulting science. In both cases, the question remains: is marketplace science trustworthy?

## The author

**Paul Lucier** is an independent historian and author of the 2008 book *Scientists and Swindlers: Consulting on Coal and Oil in America, 1820–1890*. His forthcoming book is *Science and Capitalism*.
e-mail: paullucier2@gmail.com

1. Lucier, P. *Scientists and Swindlers: Consulting on Coal and Oil in America, 1820–1890* (Johns Hopkins Univ. Press, 2008).
2. Bud, R. *Isis* **103**, 515–563 (2012).
3. Homburg, E. *Isis* **109**, 565–576 (2018).
4. Israel, P. *Edison: A Life of Invention* (Wiley, 1998).
5. Mowery, D. C. & Rosenberg, N. *Technology and the Pursuit of Economic Growth* (Cambridge Univ. Press, 1989).
6. Hounshell, D. A. & Kenly Smith Jr, J. *Science and Corporate Strategy: DuPont R&D, 1902–1980* (Cambridge Univ. Press, 1988).
7. Reich, L. S. *The Making of American Industrial Research: Science and Business at GE and Bell, 1876–1926* (Cambridge Univ. Press, 1985).
8. Usselman, S. W. in *The Oxford Encyclopedia of the History of American Science, Medicine, and Technology* (ed. Slotten, H. R.) Vol. II, 369–387 (Oxford Univ. Press, 2014).
9. Backhouse. R. E. & Maas, H. *Isis* **108**, 82–106 (2017).
10. Grandin, K., Wormbs, N. & Widmalm, S. (eds) *The Science-Industry Nexus: History, Policy, Implications* (Science History Publications, 2004).
11. Knowles, S. G. & Leslie, S. W. *Isis* **92**, 1–33 (2001).
12. Clarke, S. H., Lamoreaux, N. R. & Usselman, S. W. (eds) *The Challenge of Remaining Innovative: Insights from Twentieth Century American Business* (Stanford Business Books, 2009).
13. Landes, D. S. *The Unbound Prometheus: Technological Change and Industrial Development in Western Europe from 1750 to the Present* (Cambridge Univ. Press, 1969).
14. Hounshell, D. A. in *Engines of Innovation: U.S. Industrial Research at the End of an Era* (eds Rosenbloom, R. S. & Spencer, W. J.) 13–85 (Harvard Business School Press, 1996).
15. National Science Board. *Science and Engineering Indicators 2008* Vol. 1, 6-35–6-36 (National Science Foundation, 2008).
16. Mowery, D. C., Nelson, R. R., Sampat, B. & Ziedonis, A. *Ivory Tower and Industrial Innovation: University–Industry Technology Transfer Before and After the Bayh–Dole Act* (Stanford Univ. Press, 2004).
17. Pisano, G. P. *Science Business: The Promise, the Reality, and the Future of Biotech* (Harvard Business School Press, 2006).
18. Chesbrough, H. *Open Innovation: The New Imperative for Creating and Profiting from Technology* (Harvard Business School Press, 2003).
19. Mowery, D. C. *Ind. Corp. Change* **18**, 1–50 (2009).
20. Arora, A., Belenzon, S. & Patacconi, A. *Ind. Corp. Change* **28**, 289–307 (2019).
21. Mirowski, P. *Science-Mart: Privatizing American Science* (Harvard Univ. Press, 2011).

# Correspondence

## CAS: with the people and the government

Your Editorial marking the 70th anniversary of the Chinese Academy of Sciences (CAS) points out some of the remarkable results the organization has achieved since its inception (see *Nature* **574**, 5; 2019). However, we find your take on its history quite misleading.

For example, you seem to overplay the modern significance of China's Cultural Revolution. The devastating consequences of that have been recognized by the Chinese government for more than 40 years, and it has painstakingly implemented measures to reverse the negative effects. This great governance has ensured that China has witnessed huge advances ever since.

CAS is not run independently of government, as you imply. The establishment and development of CAS have been entirely based on the wisdom and support of the central government. The role of the academy in leading China's research has always been recognized by China's leadership, which has respected science and technology from the start — for its own sake as well as for developing a sustainable economy.

Contrary to your headline, CAS has never sought or achieved financial autonomy. Over the past 40 years, half of its income has come directly from central-government investment; the rest has been from competitive funding or technology transfer. CAS could not develop without the funding and support of the central government. And CAS is committed to facilitating technology transfer to support economic development, although it does not directly invest in the industrial sector.

The academy has a list of notable achievements, apart from those you mention. It started China's first talent programme, attracting top-quality overseas-trained scholars back to China. And CAS intends to become a leading research institution that satisfies scientific interests and regional or global needs. We have already established 10 joint research and education centres overseas and, together with another 36 science organizations, have launched the Alliance of International Science Organizations to address shared challenges and to contribute to the United Nations Sustainable Development Goals.

You suggest that CAS could be a model for science academies in other countries — particularly in one-party states or those with authoritarian leadership. Our core competence lies in our unique role as a national research institution. Although every academy should of course determine its own development, we find that an integrated structure combining research, education, consultation and technology transfer suits us well.

We object to your allegation that the Chinese central government takes "harsh measures against its people". In carrying out its scientific and technical mission, CAS stands firmly with the central government and with the people. We reject any such false allegations with disruptive intentions and are strongly opposed to biased judgments of China's internal affairs, and to any unnatural linking of political or ideological positions with our mission.

**Qingquan Zhang** Bureau of International Cooperation, Chinese Academy of Sciences, Beijing, China.
qqzhang@cashq.ac.cn

## Italy's evaluators: rankings boom is real

As president and vice-president of the Italian National Agency for the Evaluation of Universities and Research Institutes (ANVUR), we disagree that Italy has been climbing the international research-impact rankings because Italian scholars are citing each other's articles more heavily (*Nature* **572**, 578–579; 2019).

Scientific productivity in Italy has risen in the past decade, possibly stimulated by the introduction of performance-related university funding. Such systems tend to increase a country's publications in the short term, as well as to boost the number of citations per paper when normalized for each field. The use of metrics can itself have positive effects on scientific output (see D. Checchi *et al. High. Educ. Q.* **73**, 45–69; 2019).

ANVUR recognizes the importance of correcting gaming behaviour, including self-citation. In our most recent evaluation exercise (in 2011–14), papers in which self-citation exceeded a given threshold were downgraded. We intend to seek evidence of gaming behaviour at the individual and article level, and clamp down on it in future evaluations if necessary.

The Italian research system has responded to public demand for more transparency and accountability. Citation doping alone cannot explain the concomitant rise in publications and citations (see also P. D'Antuono and M. Ciavarella *Nature* **574**, 333; 2019). The rise should in fact be viewed with some pride by the Italian scientific community.

**Paolo Miccoli**, **Raffaella I. Rumiati** ANVUR, Rome, Italy.
raffaella.rumiati@anvur.it

## China's silicon valley must protect nature

We welcome the development of Xi'an — China's former capital and the original eastern end of the Silk Road — into a high-tech city at the heart of the country's 2013 Belt and Road Initiative for worldwide trade (*Nature* **563**, S25–S27; 2018). However, it is crucial that the ambitious infrastructure planning includes provisions to protect the city's environment from further degradation.

The nearby Qinling Mountains provide 90% of the drinking water for the 10 million or so residents of Xi'an. The range hosts 4,000 plant and animal species, and contains 15 natural and cultural heritage sites of ancient civilizations going back 5,000 years to the Xia dynasty. Developing China's 'silicon valley' so close to these mountains could seriously disrupt the ecosystem (S. Thacker *et al. Nature Sustain.* **2**, 324–331; 2019).

Xi'an is already one of China's most polluted cities, with many outdated coal-burning factories. Only half of the city's 15 rivers are classified as clean. The daily discharge of domestic sewage into these rivers can reach 8,000 tonnes. The Zao River, which crosses the city's Hi-tech Industries Development Zone, is black and malodorous.

All these issues need to be addressed before major changes associated with the development go ahead (see, for example, L. Han *et al. Sci. Rep.* **6**, 23604; 2016).

**Zhenhong Wang**, **He Li**, **Amit Patil** Chang'an University, Xi'an, China.
w_zhenhong@126.com

# News & views

Quantum information

# Quantum computing takes flight

William D. Oliver

A programmable quantum computer has been reported to outperform the most powerful conventional computers in a specific task – a milestone in computing comparable in importance to the Wright brothers' first flights. **See p.505**

Quantum computers promise to perform certain tasks much faster than ordinary (classical) computers. In essence, a quantum computer carefully orchestrates quantum effects (superposition, entanglement and interference) to explore a huge computational space and ultimately converge on a solution, or solutions, to a problem. If the numbers of quantum bits (qubits) and operations reach even modest levels, carrying out the same task on a state-of-the-art supercomputer becomes intractable on any reasonable timescale – a regime termed quantum computational supremacy[1]. However, reaching this regime requires a robust quantum processor, because each additional imperfect operation incessantly chips away at overall performance. It has therefore been questioned whether a sufficiently large quantum computer could ever be controlled in practice. But now, on page 505, Arute *et al.*[2] report quantum supremacy using a 53-qubit processor.

Arute and colleagues chose a task that is related to random-number generation: namely, sampling the output of a pseudo-random quantum circuit. This task is implemented by a sequence of operational cycles, each of which applies operations called gates to every qubit in an $n$-qubit processor. These operations include randomly selected single-qubit gates and prescribed two-qubit gates. The output is then determined by measuring each qubit.

The resulting strings of 0s and 1s are not uniformly distributed over all $2^n$ possibilities. Instead, they have a preferential, circuit-dependent structure – with certain strings being much more likely than others because of quantum entanglement and quantum interference. Repeating the experiment and sampling a sufficiently large number of these solutions results in a distribution of likely outcomes. Simulating this probability distribution on a classical computer using even today's leading algorithms becomes exponentially more challenging as the number of qubits and operational cycles is increased.

In their experiment, Arute *et al.* used a quantum processor dubbed Sycamore. This processor comprises 53 individually controllable qubits, 86 couplers (links between qubits) that are used to turn nearest-neighbour two-qubit interactions on or off, and a scheme to measure all of the qubits simultaneously. In addition, the authors used 277 digital-to-analog converter devices to control the processor.

When all the qubits were operated simultaneously, each single-qubit and two-qubit gate had approximately 99–99.9% fidelity – a measure of how similar an actual outcome of an operation is to the ideal outcome. The attainment of such fidelities is one of the remarkable technical achievements that enabled this work. Arute and colleagues determined the fidelities using a protocol known as cross-entropy benchmarking (XEB). This protocol was introduced last year[3] and offers certain advantages over other methods for diagnosing systematic and random errors.

The authors' demonstration of quantum supremacy involved sampling the solutions from a pseudo-random circuit implemented on Sycamore and then comparing these results to simulations performed on several powerful classical computers, including the Summit supercomputer at

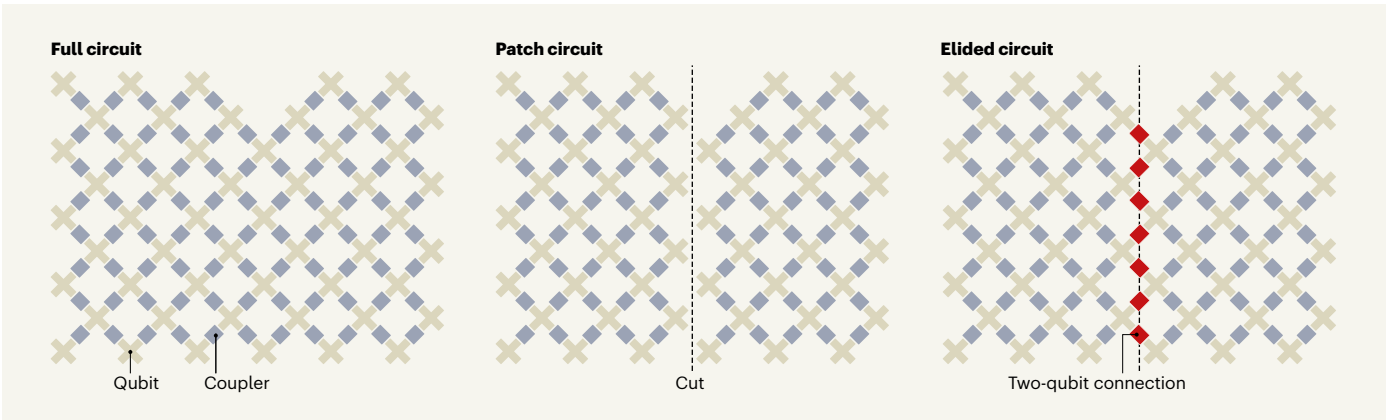



**Figure 1 | Three types of quantum circuit.** Arute *et al.*[2] demonstrate that a quantum processor containing 53 quantum bits (qubits) and 86 couplers (links between qubits) can complete a specific task much faster than an ordinary computer can simulate the same task. Their demonstration is based on three quantum circuits: the full circuit, the patch circuit and the elided circuit. The full circuit comprises all 53 qubits and is the hardest to simulate on an ordinary computer. The patch circuit cuts the full circuit into two patches that are each relatively easy to simulate. Finally, the elided circuit links these two patches using a reduced number of two-qubit operations along reintroduced two-qubit connections and is intermediate between the full and patch circuits, in terms of its ease of simulation.

Oak Ridge National Laboratory in Tennessee (see go.nature.com/35zfbuu). Summit is currently the world's leading supercomputer, capable of carrying out about 200 million billion operations per second. It comprises roughly 40,000 processor units, each of which contains billions of transistors (electronic switches), and has 250 million gigabytes of storage. Approximately 99% of Summit's resources were used to perform the classical sampling.

Verifying quantum supremacy for the sampling problem is challenging, because this is precisely the regime in which classical simulations are infeasible. To address this issue, Arute *et al.* first carried out experiments in a classically verifiable regime using three different circuits: the full circuit, the patch circuit and the elided circuit (Fig. 1). The full circuit used all $n$ qubits and was the hardest to simulate. The patch circuit cut the full circuit into two patches that each had about $n/2$ qubits and were individually much easier to simulate. Finally, the elided circuit made limited two-qubit connections between the two patches, resulting in a level of computational difficulty that is intermediate between those of the full circuit and the patch circuit.

The authors selected a simplified set of two-qubit gates and a limited number of cycles (14) to produce full, patch and elided circuits that could be simulated in a reasonable amount of time. Crucially, the classical simulations for all three circuits yielded consistent XEB fidelities for up to $n = 53$ qubits, providing evidence that the patch and elided circuits serve as good proxies for the full circuit. The simulations of the full circuit also matched calculations that were based solely on the individual fidelities of the single-qubit and two-qubit gates. This finding indicates that errors remain well described by a simple, localized model, even as the number of qubits and operations increases.

Arute and colleagues' longest, directly verifiable measurement was performed on the full circuit (containing 53 qubits) over 14 cycles. The quantum processor took one million samples in 200 seconds to reach an XEB fidelity of 0.8% (with a sensitivity limit of roughly 0.1% owing to the sampling statistics). By comparison, performing the sampling task at 0.8% fidelity on a classical computer (containing about one million processor cores) took 130 seconds, and a precise classical verification (100% fidelity) took 5 hours. Given the immense disparity in physical resources, these results already show a clear advantage of quantum hardware over its classical counterpart.

The authors then extended the circuits into the not-directly-verifiable supremacy regime. They used a broader set of two-qubit gates to spread entanglement more widely across the full 53-qubit processor and increased the number of cycles from 14 to 20. The full circuit could not be simulated or directly verified in a reasonable amount of time, so Arute *et al.* simply archived these quantum data for future reference – in case extremely efficient classical algorithms are one day discovered that would enable verification. However, the patch-circuit, elided-circuit and calculated XEB fidelities all remained in agreement. When 53 qubits were operating over 20 cycles, the XEB fidelity calculated using these proxies remained greater than 0.1%. Sycamore sampled the solutions in a mere 200 seconds, whereas classical sampling at 0.1% fidelity would take 10,000 years, and full verification would take several million years.

This demonstration of quantum supremacy over today's leading classical algorithms on the world's fastest supercomputers is truly a remarkable achievement and a milestone for quantum computing. It experimentally suggests that quantum computers represent a model of computing that is fundamentally different from that of classical computers[4]. It also further combats criticisms[5,6] about the controllability and viability of quantum computation in an extraordinarily large computational space (containing at least the $2^{53}$ states used here).

However, much work is needed before quantum computers become a practical reality. In particular, algorithms will have to be developed that can be commercialized and operate on the noisy (error-prone) intermediate-scale quantum processors that will be available in the near term[1]. And researchers will need to demonstrate robust protocols for quantum error correction that will enable sustained, fault-tolerant operation in the longer term.

Arute and colleagues' demonstration is in many ways reminiscent of the Wright brothers' first flights. Their aeroplane, the *Wright Flyer*, wasn't the first airborne vehicle to fly, and it didn't solve any pressing transport problem. Nor did it herald the widespread adoption of planes or mark the beginning of the end for other modes of transport. Instead, the event is remembered for having shown a new operational regime – the self-propelled flight of an aircraft that was heavier than air. It is what the event represented, rather than what it practically accomplished, that was paramount. And so it is with this first report of quantum computational supremacy.

**William D. Oliver** is in the Departments of Electrical Engineering and Computer Science and of Physics, Research Laboratory for Electronics and Lincoln Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.
e-mail: william.oliver@mit.edu

1. Preskill, J. Preprint at https://arxiv.org/abs/1203.5813 (2012).
2. Arute, F. *et al. Nature* **574**, 505–511 (2019).
3. Boixo, S. *et al. Nature Phys.* **14**, 595–600 (2018).
4. Bernstein, E. & Vazirani, U. *Proc. 25th Annu. Symp. Theory Comput.* (ACM, 1993).
5. Dyakonov, M. The case against quantum computing. *IEEE Spectrum* (2018).
6. Kalai, G. Preprint at https://arxiv.org/abs/1908.02499 (2019).

Neuroscience

# Gut microbes help mice forget their fear

**Drew D. Kiraly**

Microorganisms in the gut influence fear-related learning. The results of a study that reveals some of the mechanistic underpinnings of this phenomenon promise to boost our understanding of gut–brain communication. **See p.543**

The gut's resident bacteria, collectively called the gut microbiota, can have marked effects on brain function and on behaviour – but the mechanisms underlying this interplay remain largely unknown. On page 543, Chu *et al.*[1] define these mechanisms in unprecedented scope and detail. The authors report that mice lacking a complex microbiota exhibit altered fear-associated behaviour, changes in gene expression in cells in the brain, and alterations in the firing patterns and rewiring ability of neurons. The work represents a leap forward in our understanding of the interplay between the gut and brain.

Animals update their responses to environmental cues throughout their lives. This process of behavioural adaptation is driven by underlying cellular and molecular changes in the brain. Chu and colleagues analysed how changes in the gut microbiota affect one such adaptation: fear conditioning.

First, the authors trained mice to associate a tone with an electric shock, and measured how strongly that association was formed. The association developed normally both in control animals and in animals that had been treated with antibiotics to deplete their gut microbiota. The researchers then performed an extinction task, in which they repeatedly played the tone without an electric shock before measuring

the rate at which the animals updated their behaviour (such an update indicates that the fear response has been extinguished). The microbiota-deficient mice were unable to update their response, and showed persistent fearful behaviour long after control animals had adapted. Chu *et al.* found the same phenomenon in mice that had been raised germ-free in sterile isolators and so had never developed a gut microbiota.

The current study is not the first to examine the effects of the microbiota on fear conditioning — previous work has shown a decrease in the acquisition of this response in germ-free mice compared with controls[2,3]. But Chu and colleagues are the first to report a specific deficit in fear extinction (Fig. 1). What truly sets their work apart, however, is the breadth and depth of the mechanistic findings that they subsequently went on to gather.

Extinction of the fear response is heavily dependent on the function of the brain's prefrontal cortex[4]. Chu *et al.* performed *in vivo* imaging of this brain region in their animals to analyse both neuronal activity patterns and the formation and elimination of structures called dendritic spines, which are involved in the formation of synaptic connections between neurons. During the fear-extinction test, control animals showed less dendritic-spine elimination and more spine formation than did microbiota-deficient animals. The ability to create synapses and to maintain appropriate existing synapses is a key part of synaptic plasticity — a process crucial to learning and memory, in which the strength of synaptic connections changes in response to changes in neuronal activity. A higher ratio of spine formation to elimination might therefore partially explain why control animals were better able to appropriately extinguish the fearful stimulus.

Tight control of gene expression is also crucial for proper regulation of synaptic and behavioural plasticity. Previous work has indicated that changes in the microbiota alter the gene-expression profile of the prefrontal cortex as a whole[5], but Chu and colleagues performed RNA sequencing on single cells throughout the region, enabling them to identify gene-expression changes in individual cell types. These data show that microbiota depletion has a more pronounced effect on excitatory than on inhibitory neurons, setting the stage for future research in which the microbiota could be targeted to alter the characteristics of specific neuronal populations.

The authors' single-cell sequencing also reveals gene-expression changes in microglia, the brain's resident immune cells. Previous studies[6,7] have shown that altering the microbiota causes changes in microglial gene expression and function. Chu and colleagues found high expression of genes associated with an immature state in the microglia of their microbiota-deficient animals — a change that might affect the cells' ability to function normally.

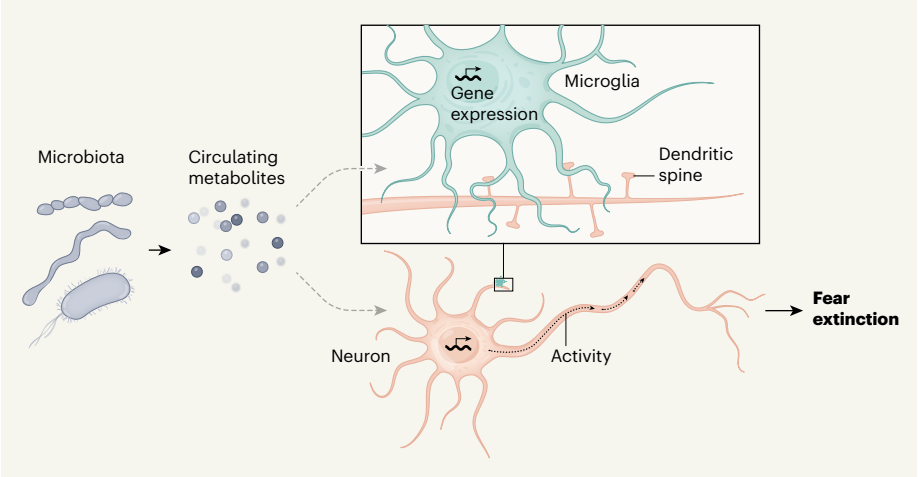



**Figure 1 | Multiple effects of the gut microbiota on the brain.** The gut's resident bacteria, the microbiota, can markedly affect the brain and behaviour. Chu *et al.*[1] provide evidence that the microbiota is needed for mice to update their behaviour in response to changing environmental cues — for example, to stop reacting to a once-frightening stimulus when it is no longer threatening (a phenomenon called fear extinction). The authors hypothesize that this role in behavioural adaptation involves metabolite molecules that are produced by the microbiota and circulate in the blood. They suggest that the metabolites modulate the ability of the brain's immune cells, microglia, to engulf and degrade structures called dendritic spines that form synaptic connections between neurons. In addition, microglia could affect neuronal activity directly — together, these activities would promote behavioural adaptation. In support of this idea, the researchers show that changes in the microbiota lead to altered gene expression in microglia and neurons, and to changes in dendritic-spine maintenance.

In the past decade, it has become clear that microglia have a crucial role in synaptic connectivity. By engulfing and degrading unwanted synapses, the cells ensure that neuronal connections are pruned or maintained as needed[8]. Changes in this process can alter neurodevelopment[9] and are implicated in psychiatric disease[10]. The researchers' RNA sequencing revealed changes in genes related to the role of the microglia in synapse organization and assembly. Although Chu *et al.* did not directly assess changes in the engulfment of synapses, their results lay the groundwork for future research into how interactions between the microbiota and microglia affect synapse density in the brain.

Finally, Chu and colleagues profiled gut metabolites (the molecules produced from metabolic processes) to identify molecules that might drive the gut–brain interactions they had observed. The authors found four metabolites that were significantly less abundant in microbiota-deficient mice than in controls. They therefore posit that the microbiota affects neurons and microglia in the brain through metabolites that are released into the circulation.

The gut microbiota is highly metabolically active, and the theory that the gut and brain communicate through circulating microbiota-derived metabolites is a popular one[11]. Manipulations of microbial metabolites have been shown to affect a range of behaviours, from autism-like actions[12] to those involving reward-seeking for drugs[13]. Experiments that manipulate levels of the metabolites identified by Chu *et al.* could improve our understanding of gut–brain communication.

Such research could also reveal a route to translating the current findings into clinical advances. The potential applications are wide-ranging, because alterations in cognition and synaptic plasticity are seen in nearly all neuropsychiatric disorders. Perhaps most germane to the current study would be the treatment of post-traumatic stress disorder, in which people cannot extinguish memories of frightening or traumatic experiences. Chu and colleagues' work raises the possibility of targeting the gut microbiota and its metabolites as a strategy for helping such individuals. Much remains to be done, but this study is an important step in our mechanistic understanding of the gut–brain axis.

**Drew D. Kiraly** is in the Department of Psychiatry and the Nash Family Department of Neuroscience, Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA.
e-mail: drew.kiraly@mssm.edu

1. Chu, C. *et al. Nature* **574**, 543–548 (2019).
2. Hoban, A. E. *et al. Mol. Psychiatry* **23**, 1134–1144 (2018).
3. Lu, J. *et al. PLoS ONE* **13**, e0201829 (2018).
4. Maren, S., Phan, K. L. & Liberzon, I. *Nature Rev. Neurosci.* **14**, 417–428 (2013).
5. Hoban, A. E. *et al. Transl. Psychiatry* **6**, e774 (2016).
6. Erny, D. *et al. Nature Neurosci.* **18**, 965–977 (2015).
7. Thion, M. S. *et al. Cell* **172**, 500–516 (2018).
8. Schafer, D. P. *et al. Neuron* **74**, 691–705 (2012).
9. Zhan, Y. *et al. Nature Neurosci.* **17**, 400–406 (2014).
10. Sekar, A. *et al. Nature* **530**, 177–183 (2016).
11. Cryan, J. F. & Dinan, T. G. *Nature Rev. Neurosci.* **13**, 701–712 (2012).
12. Hsiao, E. Y. *et al. Cell* **155**, 1451–1463 (2013).
13. Kiraly, D. D. *et al. Sci. Rep.* **6**, 35455 (2016).

Astrochemistry

# The origins of buckyballs in space

Alessandra Candian

The spectroscopic fingerprints of buckyballs have been observed in space, but questions remain about how these large molecules form. Laboratory experiments have revealed a possible mechanism.

A long-standing mystery in astronomical spectroscopy concerns diffuse interstellar bands, a family of absorption features seen in the spectra of the interstellar medium of the Milky Way and of other galaxies. First observed almost 100 years ago, the origin of any of the bands was unknown until 2015, when four of them were assigned[1] to the cation of buckminsterfullerene ($C_{60}^+$; the uncharged molecule is often referred to simply as fullerene, or colloquially as a buckyball). Fullerene and its analogue, $C_{70}$, are by far the biggest molecules detected in space, raising the question of how such large species can form in those rarified conditions. Researchers have suggested that fullerene forms in the outflows of old, carbon-rich stars known as asymptotic giant branch stars[2] — the temperatures and densities of these outflows promote chemistry similar to that of combustion. This could lead to the formation of soot, which can contain fullerene-like structures. Writing in *Astrophysical Journal Letters*, Bernal *et al.*[3] propose a very different formation route for fullerene.

The carbon atoms in fullerene are arranged in the shape of a football, a molecular structure that is remarkably stable but also difficult to construct. Fullerene has been made in the laboratory in experiments designed to probe the chemistry that occurs in carbon-rich stars: carbon in the form of graphite was vaporized into a high-density helium flow, producing carbon clusters[4]. The discovery that fullerene was among the reaction products led to the award of the Nobel Prize in Chemistry to Harry Kroto, Richard Smalley and Robert Curl in 1996.

However, the range of temperatures required to create fullerene in this way is quite specific[2]; outside that range, molecules known as polycyclic aromatic hydrocarbons (PAHs) are produced instead. These molecules are 2D sections of a single layer of graphite (a graphene sheet), decorated with hydrogen atoms. Subsequent experiments[5,6] have shown that PAHs that contain more than 60 carbon atoms are converted into fullerenes when exposed to sufficient ultraviolet irradiation.

The first astronomical source in which fullerene was detected was the star Tc1 (ref. 7). Puzzlingly, however, the emission associated with fullerene came from a location far away from the star and its ultraviolet photons, whereas the PAH emissions were closer to the star. On the basis of the previously reported laboratory experiments, this is the opposite of what should happen if fullerene forms from PAHs in this source[8]. So how can the locations of the emissions be explained?

Bernal and co-workers now report that fullerene also forms readily from silicon carbide (SiC), which has been proposed to be the first carbonaceous material to condense out of old, carbon-rich stars[9]. The authors rapidly heated grains of the crystalline form of SiC that is found in highest abundance in meteorites[10], and irradiated them with xenon ions, mimicking the heating caused by shock waves around old stars.

Using a transmission electron microscope to image the surfaces of the samples down to the subnanometre scale, the scientists observed that the grain material had altered notably as a result of its treatment (Fig. 1). Silicon atoms had percolated to the outer layers of the grains, leaving behind what looked like sheets of carbon atoms in a hexagonal 'chicken-wire' arrangement — that is, graphene sheets.

The transformation of the outer layers of SiC into graphene sheets at high temperatures had been reported[11] previously for a different form of SiC from that studied by Bernal and colleagues. However, Bernal *et al.* also observed the formation of hemispherical structures with diameters similar to that of fullerene. Their work thus provides a convincing new mechanism for the formation of fullerene in evolved stars.

Bernal *et al.* report another piece of evidence supporting the idea that SiC grains are rapidly heated and bombarded with ions in evolved stars. They have identified a fragment of the Murchison meteorite — a highly studied meteorite that is rich in organic compounds — in which the ratio of carbon-12 to carbon-13 isotopes is typical of material from an old, carbon-rich star. This indicates that the fragment was not produced during or after the formation of the meteorite, but instead is stardust that originated in an old star. The fragment has a core of SiC surrounded by graphene sheets. However, previous analyses[12] of graphite-containing stardust found evidence only of titanium carbide cores, rather than SiC cores. This raises the question of how common SiC cores are in graphite-containing stardust.

The rapid heating of SiC grains in the presence of hydrogen can lead to the formation of PAHs[13]. Bernal and colleagues' findings therefore suggest that the thermal conversion of SiC to graphene sheets in evolved stars

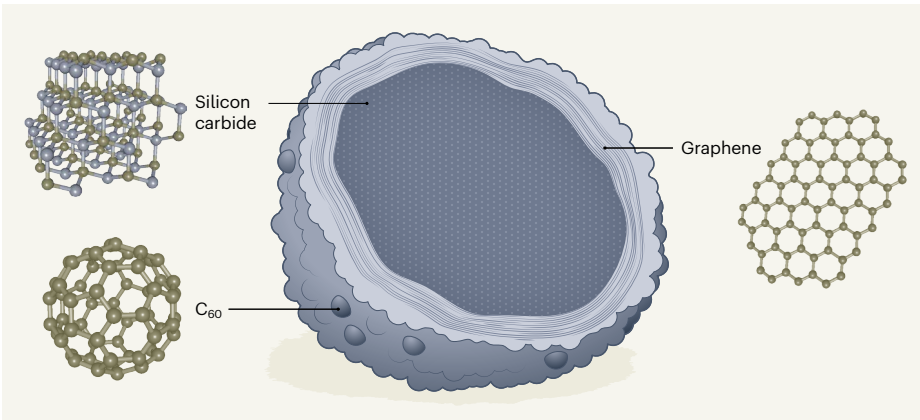



**Figure 1 | Evidence of a mechanism for the formation of buckminsterfullerene in space.** Bernal *et al.*[3] heated grains of silicon carbide (SiC) and bombarded them with ions, mimicking the conditions experienced by the dust around old stars. Using a transmission electron microscope, the scientists observed that the outer layers of SiC had transformed into graphene sheets, as shown in this idealized grain. They also observed the formation of hemispherical structures with diameters similar to that of buckminsterfullerene ($C_{60}$) on the surface of the grains. Their work thus reveals a convincing process through which $C_{60}$ could form in the outflows of old stars.

could be the first step in the formation of large carbon-containing molecules in general: subsequent (or simultaneous) exposure of the graphene to atomic hydrogen produces PAHs, whereas ion bombardment produces fullerene. Alternatively, PAH molecules might be molecular intermediates in the formation of carbon soot, which can then be broken down by ultraviolet irradiation to make PAHs again[14].

The efficiency of Bernal and colleagues' fullerene-forming mechanism is unknown, raising the question of how many SiC grains are needed to account for the observed abundance of fullerene molecules in space. If there aren't enough grains, then a further mechanism will be required to explain the abundance of fullerene. By contrast, if there are too many SiC grains, what happens to the 'excess' fullerene molecules produced, given that they are notoriously difficult to degrade? More experiments and detailed modelling of the formation of fullerene and of other carbon-containing large molecules from SiC grains are needed to understand this process, and to quantify its importance in old stars.

The launch of the James Webb Space Telescope in 2021 will provide powerful new tools for studying old stars, among other astronomical objects. Observations of fullerene-containing sources[7,8] such as Tc 1 will be able to constrain the regions in which SiC grains, fullerene and PAHs are present, providing more clues about how large molecules are actually formed. Further analysis and modelling of the routes involved will eventually allow astronomers to suggest the identities of the other mysterious molecules responsible for the diffuse interstellar bands.

**Alessandra Candian** is at the van't Hoff Institute for Molecular Sciences, University of Amsterdam, 1090 GD Amsterdam, the Netherlands, and at the Leiden Observatory, University of Leiden, Leiden, the Netherlands.
e-mail: a.candian@uva.nl

1. Campbell, E. K., Holz, M., Gerlich, D. & Maier, J. P. *Nature* **523**, 322–323 (2015).
2. Jäger, C., Huisken, F., Mutschke, H., Llamas Jansa, I. & Henning, T. *Astrophys. J.* **696**, 706–712 (2009).
3. Bernal, J. J. *et al. Astrophys. J.* **883**, L43 (2019).
4. Kroto, H. W., Heath, J. R., O'Brien, S. C., Curl, R. F. & Smalley, R. E. *Nature* **318**, 162–163 (1985).
5. Berné, O. & Tielens, A. G. G. M. *Proc. Natl Acad. Sci. USA* **109**, 401–406 (2012).
6. Zhen, J., Castellanos, P., Paardekooper, D. M., Linnartz, H. & Tielens, A. G. G. M. *Astrophys. J.* **797**, L30 (2014).
7. Cami, J., Bernard-Salas, J., Peeters, E. & Male, S. E. *Science* **329**, 1180–1182 (2010).
8. Cami, J., Peeters, E., Bernard-Salas, J., Doppmann, G. & De Buizer, J. *Galaxies* **6**, 101 (2018).
9. Frenklach, M., Carmer, C. S. & Feigelson, E. D. *Nature* **339**, 196–198 (1989).
10. Daulton, T. L. *et al. Science* **296**, 1852–1855 (2002).
11. Mishra, N., Boeckl, J., Motta, N. & Iacopi, F. *Phys. Status Solidi A* **213**, 2277–2289 (2016).
12. Croat, K. T., Bernatowicz, T. J. & Daulton, T. L. *Elements* **10**, 441–446 (2014).
13. Merino, P. *et al. Nature Commun.* **5**, 3054 (2014).
14. Cherchneff, I., Barker, J. R. & Tielens, A. G. G. M. *Astrophys. J.* **401**, 269–287 (1992).

Optical physics

# Light trapping gets a boost

**Kirill Koshelev & Yuri Kivshar**

The ability of structures called optical resonators to trap light is often limited by scattering of light off fabrication defects. A physical mechanism that suppresses this scattering has been reported that could lead to improved optical devices. **See p.501**

Devices called optical resonators confine light, but for only a limited time because of unavoidable light emission. On page 501, Jin *et al.*[1] report that such emission can be greatly reduced by using the interference of light waves known as bound states in the continuum. Such waves are akin to exotic electron waves that were introduced in the theory of quantum mechanics almost a century ago[2]. The authors' finding could have many technological implications for nanophotonics, quantum optics and nonlinear optics — the study of how intense light interacts with matter.

Interference is a common wave phenomenon in physics, whereby two or more waves pass through one another to produce a combined waveform. Consider the case in which these waves are correlated with one another, either because they come from the same source or because they have almost the same frequency. If the crest of one wave coincides with the crest of another wave, the combined amplitude will be the sum of the individual amplitudes. And if the crest of one wave meets the trough of another wave, the combined amplitude will be the difference in the individual amplitudes. These two scenarios are called constructive and destructive interference, respectively.

The effects of interference can be observed for all waves, but interference associated with bound states in the continuum (BICs) has attracted much attention in photonics over the past few years[3]. BICs are formed by the destructive interference of several ordinary light waves that have a similar wavevector — a quantity that describes a wave's velocity and direction of propagation. This interference provides a means of achieving strong confinement of light and of increasing its amplitude through a phenomenon known as optical resonance. It can also be used to tune an optical resonator into the 'supercavity' regime, in which emission of light from the resonator is restrained[4]. Several approaches to realizing BICs have been suggested for waves in electronic, electromagnetic and acoustic systems.

The concept of BICs was proposed for unusual states of electron waves by two pioneers of quantum mechanics, John von Neumann and Eugene Wigner[2]. They discovered that specific potentials (potential-energy profiles) could support spatially localized electron states that have energies larger than the maximum energy of the potential. In other words, the states could be confined even though their energies would normally allow them to escape. In photonics, a light wave that is trapped by an optical resonator can be converted to a BIC under certain conditions[3] — a discovery that was made only in 2008.

The main characteristic of an optical resonator is the quality factor — the ratio of the time over which the device can trap light to the period of the wave's oscillation. If the light waves destructively interfere to form BICs, the quality factor greatly increases. Moreover, in the BIC regime, the quality factor theoretically tends to infinity when one of the system parameters, such as the size of the resonator, is tuned. By contrast, the quality factor of a

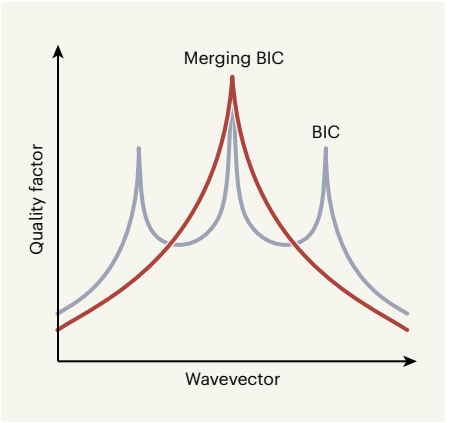



**Figure 1 | Increasing the quality factor of an optical resonator.** Jin *et al.*[1] report simulations of and experiments on a light-trapping device known as an optical resonator. The key characteristic of a resonator is the quality factor — a measure of the efficiency of light trapping. This quantity varies with the wavevector, which describes the velocity and propagation direction of a wave. The authors used their resonator to trap light in the form of waves called bound states in the continuum (BICs). They then combined these BICs into a single state: a merging BIC. As this graph shows, a merging BIC increases the quality factors of all waves that have similar wavevectors to it.

conventional resonance is not substantially affected by parameter variations.

In practical optical resonators, the quality factors of BICs are fundamentally limited by inevitable fabrication defects, which scatter light out of the plane of the device. Any light wave that is scattered off a structural imperfection changes its wavevector. To prevent scattering losses, waves must remain trapped in the resonator even after these changes have occurred. In other words, the quality factor needs to be high both before and after scattering.

Jin and colleagues have suggested and demonstrated an innovative physical mechanism for achieving optical resonances that are extremely robust to out-of-plane scattering. They considered a structure called a photonic crystal slab, consisting of a submicrometre-thick dielectric (electrically insulating) membrane patterned with a square lattice of circular holes.

The authors first ran numerical simulations to study the optical resonances in their membrane. By carefully selecting the membrane's parameters, they achieved several simulated BICs that had different wavevectors. They then altered the periodicity of the lattice until the BICs had the same wavevector. This gave rise to a new type of optical resonance: a merging BIC (which one might refer to as a super-BIC; Fig. 1). The hallmark of a merging BIC is that it increases the quality factor of all waves that have nearly the same wavevector as the resonance, reducing scattering losses from the resonator.

Jin *et al.* then experimentally demonstrated their mechanism by fabricating a set of silicon membranes that had different lattice periodicities. Some of these membranes supported a merging BIC at telecommunication wavelengths (about 1,550 nanometres) and others were close to this merging-BIC regime. The authors used a tunable telecommunication-wavelength laser to measure the intensity of scattered light along different directions for each of the samples. They found that the membranes supporting a merging BIC had a quality factor that was about 10 times larger than that for the membranes not in the merging-BIC regime. Moreover, they showed that the observed increase in quality factor was robust by finding a similar level of enhancement in all of the fabricated samples that had a merging-BIC design.

The demonstration could have many consequences for engineering high-quality resonances in nanophotonics. The ability to convert light waves into BICs allows the realization of the supercavity regime, in which highly compact resonators can have extremely large quality factors[5]. Dielectric materials that have high refractive indices could be used to reduce the resonator dimensions and to combine individual BIC resonators that have high-quality resonances into structured arrays[6].

We predict that an electromagnetic theory will be developed for describing high-quality resonances in individual dielectric nanoparticles of high refractive index and arrays of such nanoparticles, and that they all will be expressed in terms of the mathematics used to study interference in quantum mechanics. In the real world, the engineering of quality factors in the BIC regime could lead to substantial enhancement of nonlinear and quantum effects, the development of lasers that consume little power, and the realization of nanoscale resonators that facilitate strong confinement of light and large boosts to its amplitude.

**Kirill Koshelev** and **Yuri Kivshar** are at the Nonlinear Physics Centre, Australian National University, Canberra, ACT 2601, Australia, and in the Department of Nanophotonics and Metamaterials, ITMO University, St Petersburg, Russia.
e-mails: kirill.koshelev@anu.edu.au; ysk@physics.anu.edu.au

1. Jin, J. *et al. Nature* **574**, 501–504 (2019).
2. von Neumann, J. & Wigner, E. P. *Phys. Z.* **30**, 465–467 (1929).
3. Hsu, C. W., Zhen, B., Stone, A. D., Joannopoulos, J. D. & Soljačić, M. *Nature Rev. Mater.* **1**, 16048 (2016).
4. Rybin, M. & Kivshar, Y. *Nature* **541**, 165–166 (2017).
5. Rybin, M. *et al. Phys. Rev. Lett.* **119**, 243901 (2017).
6. Koshelev, K., Lepeshov, S., Liu, M., Bogdanov, A. & Kivshar, Y. *Phys. Rev. Lett.* **121**, 193903 (2018).

Epigenetics

# Lactate links metabolism to genes

**Luke T. Izzo & Kathryn E. Wellen**

Cells regulate gene expression in part through the chemical labelling of histone proteins. Discovery of a label derived from lactate molecules reveals a way in which cells link gene expression to nutrient metabolism. **See p.575**

Cellular metabolism involves the uptake, release and biochemical interconversion of nutrients to produce energy and synthesize complex molecules. The intermediates and end products of metabolism also have essential signalling functions, modulating cell signalling and gene expression in accordance with nutritional resources[1,2]. One way in which these metabolites signal is through the chemical modification of proteins such as histones. On page 575, Zhang and colleagues[3] describe their discovery of a previously unknown histone modification, lactylation, derived from the cellular metabolite lactate.

Histones are central components of chromatin — a complex of DNA and proteins that organizes and regulates the genome. They can be altered by cellular enzymes, which add chemical tags such as methyl, acetyl and phosphate groups; these epigenetic modifications to the genome affect processes such as gene expression and DNA replication and repair. Zhang *et al.* predicted that histones might also be altered by the addition of lactyl groups, and they began their search for lactylation by using a technique called mass spectrometry, which has enabled the identification of numerous protein modifications in the past few years[4]. By looking for shifts in the masses of amino-acid residues that make up histone tails, the authors deduced the presence of a modified lysine amino-acid residue, consistent with the addition of a lactyl group. Zhang *et al.* validated this finding by comparing synthetic peptides that had been chemically modified in this way with the corresponding peptides identified in cells.

The authors also used metabolic tracing with a form of lactate labelled with a stable isotope of carbon ($^{13}C_3$-lactate) to demonstrate that lactate is involved in histone lactylation. They further found that levels of lysine lactylation rose when cells were treated with increasing doses of lactate. So, histone lactylation is derived from lactate and is sensitive to lactate levels.

Lactate is an abundant metabolite produced during glycolysis — a central metabolic process in which glucose consumed by cells is broken down to generate energy. During glycolysis, glucose is converted into two pyruvate molecules; these can be either funnelled into lactate production or transported into the cellular power generators (the mitochondria), forming the intermediate acetyl coenzyme A (acetyl-CoA) and thence entering the Krebs cycle for energy production. Lactate is produced through glycolysis in various cell types, including cancer cells and immune cells. Its production is also enhanced under certain conditions, such as hypoxia (low oxygen levels), which suppresses pyruvate entry into the Krebs cycle. Zhang and colleagues' discovery that lactate is used for histone modification is intriguing both because of the metabolite's abundance and because its production, uptake and use are all subject to dynamic regulation[5].

One substantial question that the authors

aimed to address is whether lysine lactylation responds to metabolic alterations in cells. Other metabolite-derived protein modifications — such as lysine acetylation (derived from acetyl-CoA) — are metabolically sensitive[2], providing a precedent for this idea. Zhang *et al.* found that the amount of glucose available to cells grown *in vitro* dynamically regulates the lysine lactylation of histones in those cells. Furthermore, tracing of isotopically labelled glucose ($^{13}C_6$-glucose) showed that lysine lactylation depends on glycolysis. The authors used several perturbations to promote lactate production (including hypoxia and inhibitors of mitochondrial metabolism) and to suppress it (using inhibitors of pyruvate conversion to lactate). The cumulative data indicate that lysine lactylation is highly sensitive to lactate production through glycolysis.

Zhang *et al.* next sought to investigate the biological functions of lactylation, selecting macrophages as their model. Macrophages are immune cells that can take on pro-inflammatory (termed M1) or anti-inflammatory (M2) characteristics; they undergo metabolic changes that correspond to these functions[6]. For example, macrophages that encounter signs of bacterial infection activate inflammatory genes and upregulate glycolysis[6]. Zhang *et al.* stimulated macrophages with bacteria or with the bacterial component lipopolysaccharide (LPS) to induce M1 characteristics. They found that glycolysis increased and that intracellular levels of lactate rose progressively, parallelling an increase in histone lactylation (Fig. 1a). Notably, inflammatory genes typically associated with M1 characteristics were upregulated rapidly on exposure to LPS, but did not correlate with lactate levels or lysine lactylation (Fig. 1b). Instead, the increase in lysine lactylation was slower and correlated with the upregulation of homeostatic genes (those involved in maintaining a biological steady state).

The authors went on to investigate where lysine lactylation occurred in the M1 genome, as well as how the modification altered gene expression. They found that lysine lactylation was high in gene promoter regions (which mark the start points of gene transcription), and associated positively with the levels of messenger RNA produced from those genes. The authors also compared lysine lactylation and acetylation, finding lactylation at many genes that lack acetylation — suggesting distinct roles for the two modifications. Moreover, macrophages that could not produce lactate could increase the expression of inflammatory genes in response to stimulation with LPS, but could not upregulate lysine lactylation or the associated expression of homeostatic genes at later times. These temporal dynamics led the authors to propose that a delayed 'lactate timer' involving histone lactylation drives the activation of genes involved in resolving infections to help re-establish tissue homeostasis (Fig. 1b).

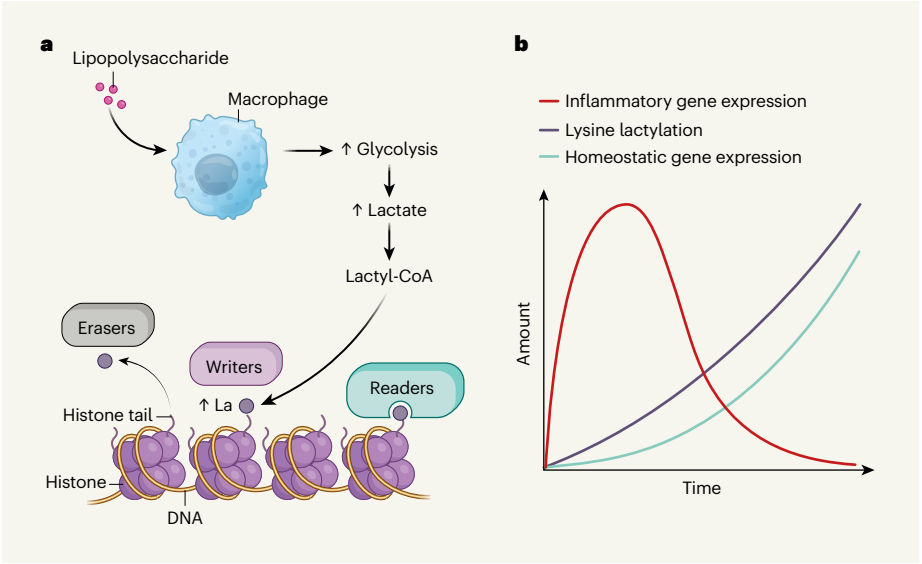



**Figure 1 | A new epigenetic modification, histone lactylation.** Zhang and colleagues[3] have discovered a chemical modification called lactylation — the addition of a lactyl (La) group to the lysine amino-acid residues in the tails of histone proteins. **a**, Stimulating immune macrophage cells with lipopolysaccharide molecules (mimicking bacterial infection) increases the conversion of glucose to energy through glycolysis. This, in turn, leads to increases in intracellular levels of the molecule lactate, and to lactylation of histones at promoter DNA sequences. However, it is unclear which enzymes generate the intermediate molecule lactyl-CoA, from which La is derived, or which enzymes deposit (writers), remove (erasers) or recognize and interpret (readers) histone lactylation. **b**, Zhang *et al.* report that the increase in lysine lactylation is delayed following macrophage stimulation. This delay correlates with changes in the expression of homeostatic genes involved in maintaining a biological steady state, but not with changes in inflammatory-gene expression. The authors therefore hypothesize that lactylation generates a 'lactate timer' to restore normal tissue function after infection.

These findings raise questions about the biochemistry of lactylation and its broader roles in physiology and disease. In terms of biochemistry, the authors show in a cell-free system that lactyl-CoA is a lactyl-group donor for lysine lactylation. So far, however, the enzymes that produce lactyl-CoA from lactate in the cell, as well as the cellular concentrations of lactyl-CoA, are unknown. Other unresolved questions concern the way in which lysine lactylation is regulated by the enzymes that deposit, read or remove this label. In the authors' cell-free system, an acetyltransferase enzyme known as p300 can catalyse the transfer of lactyl from lactyl-CoA to histones, but whether it does this in cells has yet to be tested.

**"The findings raise questions about the biochemistry of lactylation and its broader roles in physiology and disease."**

In terms of the broader roles of this modification, lactate is generated by cells both in physiological contexts — such as in skeletal muscle during exercise — and in the context of diseases such as cancer. In addition, lactate is taken up by cells of healthy tissues and tumours to feed the Krebs cycle[7]. So, as well as occurring in glycolytic cells, lactylation might also participate in communication between cells. On this note, high lactate levels in the environment around tumours are known to promote immunosuppression[8], and Zhang *et al.* found that histone lysine lactylation was greater in tumour-associated macrophages than in those from another tissue. All in all, the authors' discovery of histone lactylation provides a launch point for a deeper investigation into the roles and regulation of this modification, which links cellular metabolism to gene regulation and could have numerous implications for human health.

**Luke T. Izzo** and **Kathryn E. Wellen** are in the Department of Cancer Biology and Abramson Family Cancer Research Institute, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA.
e-mail: wellenk@upenn.edu

1. Zhu, J. & Thompson, C. B. *Nature Rev. Mol. Cell Biol.* **20**, 436–450 (2019).
2. Campbell, S. L. & Wellen, K. E. *Mol. Cell* **71**, 398–408 (2018).
3. Zhang, D. *et al. Nature* **574**, 575–580 (2019).
4. Sabari, B. R., Zhang, D., Allis, C. D. & Zhao, Y. *Nature Rev. Mol. Cell Biol.* **18**, 90–101 (2017).
5. Haas, R. *et al. Trends Biochem. Sci.* **41**, 460–471 (2016).
6. Galván-Peña, S. & O'Neill, L. A. J. *Front. Immunol.* **5**, 420 (2014).
7. Martínez-Reyes, I. & Chandel, N. S. *Cell Metab.* **26**, 803–804 (2017).
8. Ippolito, L., Morandi, A., Giannoni, E. & Chiarugi, P. *Trends Biochem. Sci.* **44**, 153–166 (2019).

Evolution

# Early tetrapods had an eye on the land

Nadia B. Fröbisch & Florian Witzmann

Fossil finds that can provide clues about how aquatic vertebrates evolved into land dwellers are elusive. But the ancient bones of a newly discovered species of tetrapod now provide some crucial missing evidence. See p.527

Following the scientific investigations into how vertebrates transitioned from water to land is like reading a good crime novel. We have a range of suspects, patchy evidence and a lot of unanswered questions. And to complicate matters, this transition from finned fish to four-limbed creatures (tetrapods) is a 'cold case' from nearly 400 million years ago. On page 527, Beznosov *et al.*[1] present some compelling detective work that sheds light on this.

The earliest-known tetrapod specimens are 380-million-year-old bone fragments that, although identifiable as belonging to a tetrapod, do not provide many details about what these animals looked like or how they lived[2]. There are also fossilized tetrapod footprints that pre-date these fossil finds by more than 14 million years[3], indicating the presence of a four-limbed, still fully aquatic track maker – but they do not reveal what the track maker looked like above the soles of its feet.

More-detailed insights into the body shape, life and growth of our early vertebrate ancestors are provided by more-complete fossil finds, including the iconic tetrapods *Acanthostega* and *Ichthyostega*[2,4]. However, these lived 365 million years ago, when tetrapods had already achieved an impressive geographical distribution and a diverse variety of body shapes and ways of life[2].

> **"This water dweller was looking above the surface of the water."**

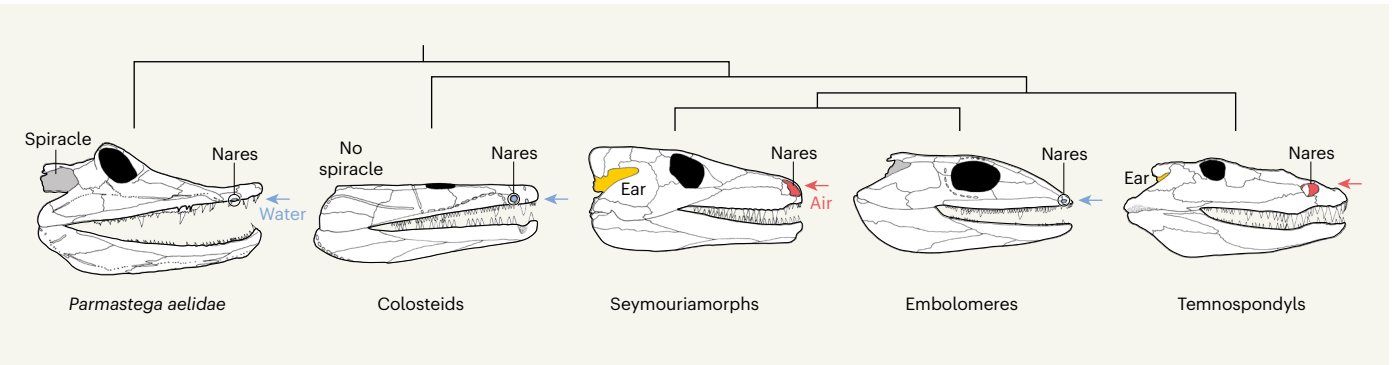
By contrast, the earliest phase of tetrapod evolution and diversification has long been mysterious. However, Beznosov and colleagues now describe skeletal fossils of a species they call *Parmastega aelidae*, which is the most ancestral (basal-most) tetrapod reported so far.

Like its known younger relatives, *P. aelidae* was a gill-breathing water dweller, and the authors estimate that this animal reached a size of more than one metre long. It lived about 372 million years ago during the Devonian period, and inhabited a shallow lagoon in a landmass that is now part of northwestern Russia. These excellently preserved fossils provide crucial data about how the major changes in breathing, sensory perception, locomotion and feeding might have taken place as tetrapods transitioned to life on land. The discovery also raises many exciting questions.

The most striking features of the *P. aelidae* skull are the large, oval-shaped eye openings, which face to the front and side, and which are positioned high up, towards the top of the skull (Fig. 1). This eye shape and position is surprising because it indicates that this water dweller was looking above the surface of the water.

Mudskippers (species from the family Oxudercidae) are modern amphibious fish that inhabit marine mud flats, and they are useful living creatures with which to compare *P. aelidae* because their eyes have a similar shape and position. Mudskippers peek above the water surface to look out for prey and potential danger[5]. But what was *P. aelidae* looking for? The need to detect enemies on land or in the air can be ruled out, because during the late Devonian period, such animals were not yet present there.

One possibility is that *P. aelidae* was looking for prey on the shore. If so, what kind of terrestrial or semi-terrestrial prey was it watching? Some have suggested that early water-dwelling tetrapods and their closest fish-like relatives might have preyed on terrestrial invertebrates of the phylum Arthropoda, which includes insects[6]. However, the large arthropods that could have provided sufficient food to sustain an animal the size of *P. aelidae* were still rare




NADIA FRÖBISCH & FLORIAN WITZMANN

**Figure 1 | The evolution of tetrapod skulls.** Beznosov *et al.*[1] report 372-million-year-old fossils of a four-limbed vertebrate (tetrapod) from just before the time when tetrapods moved onto land. They call this newly discovered species *Parmastega aelidae*. Its nasal passages (nares) are close to its jaw and would have been positioned under water. Water passing though the nares (blue arrow) would have been used for breathing when it reached the gills (not shown). *P. aelidae* could also breathe air directly through a skull opening called a spiracle (grey). Comparing these ancestral features of *P. aelidae* with other tetrapods reveals patterns of evolutionary change. The other tetrapods shown are: an early tetrapod group called colosteids; seymouriamorphs and embolomeres, members of a lineage that gave rise to amniotes (birds, reptiles and mammals); and temnospondyls, which gave rise to modern amphibians (such as frogs and salamanders). Colosteids lacked spiracles and breathed solely through their gills using water taken up through the nares. Compared with *P. aelidae*, seymouriamorphs and temnospondyls had larger and higher nares, which they would have used to breathe air (red arrow). These tetrapods lacked spiracles, and had ears (yellow) instead in that area of their skull. Embolomeres retained the breathing system used by *P. aelidae*.

in the Devonian period[7]. Moreover, *P. aelidae* had large fangs, which suggests that it preyed mainly on vertebrates. Perhaps it searched for fish carcasses stranded on the shore. Or, to make an even more speculative suggestion, maybe it scavenged early amphibious tetrapods that rested near the water. However, evidence for such creatures has not yet been found among the fossils of the Sosnogorsk Formation (the rock layers that contained the *P. aelidae* fossil).

Another notable feature of *P. aelidae* is the extremely low position, close to its jaws, of the external openings of its nose (the nares), which would have been under water (Fig. 1). This is in striking contrast to the high position of its eyes and is quite different from the configuration of nares in modern-day aquatic tetrapod animals, such as crocodiles, hippopotamuses or frogs. The eyes of those animals sit on top of their head, and their nares are likewise positioned high on the snout, which enables them to breathe air while looking above water. Judging from their submerged position, *P. aelidae* nares acted as openings through which an inflow of water was directed towards the gills during breathing. *P. aelidae* also had the option of breathing air through a large opening in its skull called a spiracle (Fig. 1), and such a breathing process would probably have been similar to that used by modern air-breathing fish[8].

This low position of the nares is found in most known early tetrapods (called stem tetrapods) of the Devonian period (approximately 419.2 million to 358.9 million years ago) and Carboniferous period (358.9 million to 298.9 million years ago). In all of these animals, the passage from the nares to the mouth cavity might still have served to transport water rather than air. Some fossils of stem tetrapods, such as those of a grouping called colosteids (Fig. 1), had lost their spiracle opening — they must therefore have relied on gill breathing. In some other early tetrapods that arose later than *P. aelidae* and were more evolved than their ancestors (a state described as being more derived), the spiracle is absent, and its place is taken by an ear[2]. These tetrapods' nares are larger and higher on the snout (Fig. 1) compared with the ancestral form, suggesting that they used their nares to transport air towards the lungs while peeking out of the water when on the lookout for prey.

The *P. aelidae* fossils offer a treasure trove of information that could help to disentangle some of the complex evolutionary changes that took place when vertebrates made the transition from aquatic to terrestrial life. This discovery also reminds us that much still remains to be learnt in the next gripping chapter of this detective story.

**Nadia B. Fröbisch** and **Florian Witzmann** are at the Natural History Museum, Leibniz Institute for Evolution and Biodiversity Science, 10115 Berlin, Germany.
e-mails: nadia.froebisch@mfn.berlin; florian.witzmann@mfn.berlin

1. Beznosov, P. A., Clack, J. A., Lukševičs, E., Ruta, M. & Ahlberg, P. E. *Nature* **574**, 527–531 (2019).
2. Clack, J. A. *Gaining Ground: The Origin and Evolution of Tetrapods* 2nd edn (Indiana Univ. Press, 2012).
3. Niedźwiedzki, G., Szrek, P., Narkiewicz, K., Narkiewicz, M. & Ahlberg, P. E. *Nature* **463**, 43–48 (2010).
4. Sanchez, S., Tafforeau, P., Clack, J. A. & Ahlberg, P. E. *Nature* **537**, 408–411 (2016).
5. Schultze, H.-P. *Sber Ges. Naturf. Freunde Berl.* **36**, 59–77 (1997).
6. MacIver, M. A., Schmitz, L., Mugan, U., Murphey, T. D. & Mobley, C. D. *Proc. Natl Acad. Sci. USA* **114**, E2375–E2384 (2017).
7. Shear, W. A. & Kukalová-Peck, J. *Can. J. Zool.* **68**, 1807–1834 (1990).
8. Graham, J. B. *et al. Nature Commun.* **5**, 3022 (2014).

Article

# Identification of strontium in the merger of two neutron stars

https://doi.org/10.1038/s41586-019-1676-3

Received: 8 February 2018

Accepted: 14 August 2019

Published online: 23 October 2019

Darach Watson[1,2*], Camilla J. Hansen[1,3,20], Jonatan Selsing[1,2,20], Andreas Koch[4], Daniele B. Malesani[1,2,5], Anja C. Andersen[1], Johan P. U. Fynbo[1,2], Almudena Arcones[6,7], Andreas Bauswein[7,8], Stefano Covino[9], Aniello Grado[10], Kasper E. Heintz[1,2,11], Leslie Hunt[12], Chryssa Kouveliotou[13,14], Giorgos Leloudas[1,5], Andrew J. Levan[15,16], Paolo Mazzali[17,18] & Elena Pian[19]

Half of all of the elements in the Universe that are heavier than iron were created by rapid neutron capture. The theory underlying this astrophysical r-process was worked out six decades ago, and requires an enormous neutron flux to make the bulk of the elements[1]. Where this happens is still debated[2]. A key piece of evidence would be the discovery of freshly synthesized r-process elements in an astrophysical site. Existing models[3–5] and circumstantial evidence[6] point to neutron-star mergers as a probable r-process site; the optical/infrared transient known as a 'kilonova' that emerges in the days after a merger is a likely place to detect the spectral signatures of newly created neutron-capture elements[7–9]. The kilonova AT2017gfo—which was found following the discovery of the neutron-star merger GW170817 by gravitational-wave detectors[10]—was the first kilonova for which detailed spectra were recorded. When these spectra were first reported[11,12], it was argued that they were broadly consistent with an outflow of radioactive heavy elements; however, there was no robust identification of any one element. Here we report the identification of the neutron-capture element strontium in a reanalysis of these spectra. The detection of a neutron-capture element associated with the collision of two extreme-density stars establishes the origin of r-process elements in neutron-star mergers, and shows that neutron stars are made of neutron-rich matter[13].

The most detailed information yet available for a kilonova comes from a series of spectra of AT2017gfo taken over several weeks with the medium-resolution, ultraviolet (320 nm) to near-infrared (2,480 nm) spectrograph X-shooter, mounted at the Very Large Telescope at the European Southern Observatory. These spectra[11,12] allow us to track the evolution of the kilonova's primary electromagnetic output from 1.5 days until 10 days after the event. Detailed modelling of these spectra has yet to be done, owing to limited understanding of the phenomenon and the expectation that a very large number of moderate to weak lanthanide lines with unknown oscillator strengths would dominate the spectra[14,15]. Despite this expected complexity we sought to identify individual elements in the early spectra, because these spectra are well reproduced by relatively simple models[11].

The first-epoch spectrum can be reproduced over the entire observed spectral range by using a single-temperature blackbody with an observed temperature of approximately 4,800 K. The two major deviations short of 1 μm from a pure blackbody are due to two very broad absorption components (with widths of roughly 0.2$c$, where $c$ is the speed of light). These components are centred at about 350 nm and 810 nm (Fig. 1). The shape of the ultraviolet absorption component is not well constrained because it lies close to the edge of our sensitivity limit and may simply be cut off below about 350 nm. The presence of the absorption feature at 810 nm in this epoch has been noted previously[11,12].

The fact that the spectrum is very well reproduced by a single-temperature blackbody in the first epoch suggests a population of states close to local thermal equilibrium (LTE). We therefore use three separate methods of increasing complexity first to determine, without too many assumptions, the most likely origin of the spectral features, and then to self-consistently model and test our conclusion. These three methods are: first, our own LTE spectral-synthesis code; second, the LTE line-analysis and spectrum-synthesis code MOOG[16]; and third, the moving-plasma radiative-transfer code TARDIS[17] (see Methods). We use a variety

[1]Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark. [2]Cosmic Dawn Center (DAWN), Copenhagen, Denmark. [3]Max-Planck-Institut für Astronomie, Heidelberg, Germany. [4]Zentrum für Astronomie der Universität Heidelberg, Astronomisches Rechen-Institut, Heidelberg, Germany. [5]DTU Space, National Space Institute, Technical University of Denmark, Kongens Lyngby, Denmark. [6]Institut für Kernphysik, Technische Universität Darmstadt, Darmstadt, Germany. [7]GSI Helmholtzzentrum für Schwerionenforschung GmbH, Darmstadt, Germany. [8]Heidelberger Institut für Theoretische Studien, Heidelberg, Germany. [9]Istituto Nazionale di Astrofisica (INAF)/Brera Astronomical Observatory, Merate, Lecco, Italy. [10]INAF/Osservatorio Astronomico di Capodimonte (OACN), Naples, Italy. [11]Centre for Astrophysics and Cosmology, Science Institute, University of Iceland, Reykjavík, Iceland. [12]INAF/Osservatorio Astrofisico di Arcetri, Firenze, Italy. [13]Department of Physics, The George Washington University, Washington, DC, USA. [14]Astronomy, Physics and Statistics Institute of Sciences (APSIS), Washington, DC, USA. [15]Department of Physics, University of Warwick, Coventry, UK. [16]Department of Astrophysics/IMAPP, Radboud University Nijmegen, Nijmegen, The Netherlands. [17]Astrophysics Research Institute, Liverpool John Moores University, Liverpool, UK. [18]Max-Planck Institute for Astrophysics, Garching, Germany. [19]INAF, Astrophysics and Space Science Observatory, Bologna, Italy. [20]These authors contributed equally: Camilla J. Hansen, Jonatan Selsing. *e-mail: darach@nbi.ku.dk

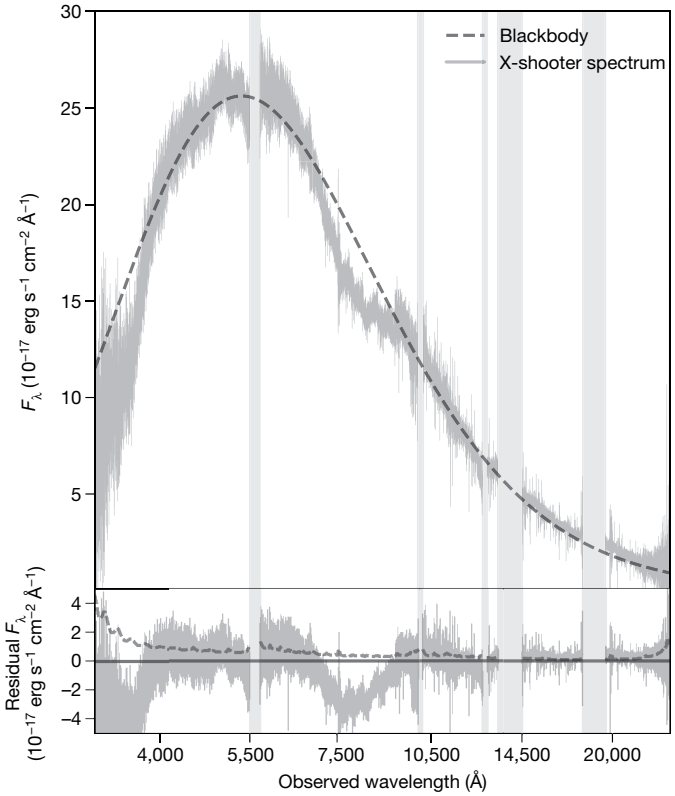



**Fig. 1 | Spectrum of the kilonova AT2017gfo, showing broad absorption features.** The spectrum shown was taken with the spectrograph X-shooter 1.5 days after the neutron-star merger GW170817. The dashed black line in the upper panel is the blackbody component of a blackbody model with broad absorption lines (see main text). The residuals of data minus blackbody are shown in the lower panel, with the dashed grey line indicating the 1$\sigma$ uncertainty on each spectral bin. The data in the sections overplotted with grey bars are affected by telluric features or are poorly calibrated regions and are not included in the fit. $F_\lambda$ is the flux per unit wavelength.

of spectral-line lists for these codes, all of which yield consistent results. For our own LTE code, we adopt a fiducial temperature of 3,700 K, which is our final model's best-fit temperature corrected by the Doppler factor (−0.23) of the absorption features that we determine below; changing the temperature of our LTE model in the range 3,700–5,100 K does not markedly affect our results.

To identify the absorption features, we seek lines with wavelengths blueshifted by 0.1–0.3$c$, corresponding approximately to 390–500 nm and 900–1,160 nm in the rest frame (see Methods). The lines will also be broadened with an observed width that depends on the velocity and geometry. For spherically expanding ejecta, the line broadening will be similar to the expansion velocity of the gas. We do not attempt a detailed geometric model here because it depends on assumptions about the geometry of the gas and the wavelength-dependent opacity, with substantial relativistic and time-delay corrections.

We adopt an initially agnostic view of the expected abundances. We use solar r-process abundance ratios (the total solar abundances of heavy elements[18] with s-process elements subtracted[19]), as well as abundances from two metal-poor stars that are old enough to be dominated by the r-process in their neutron-capture abundances[20,21]. These three sets span a wide range in their ratios of light-to-heavy r-process abundances (Fig. 2). We also produce absorption spectra for each element individually (Extended Data Figs. 1, 2).

Our LTE models using abundances from a solar-scaled r-process and metal-poor stars all show that Sr produces a strong feature centred at an observed wavelength of roughly 800 nm, as well as features at wavelengths shorter than around 400 nm, for our adopted blueshift (Fig. 3; see also Extended Data Fig. 3). The restframe wavelengths of the longer-wavelength features are 1,000–1,100 nm. It is worth noting that Sr is typically considered an s-process element because only about 30% of the cosmic (solar) abundance is produced by the r-process[18,19]. For this reason it has not always been considered in kilonova simulations. However, it is one of the more abundant r-process elements, accounting for at least a few per cent by mass of all such elements[19]. Of all of the r-process elements, Sr displays by far the strongest absorption features in this region of the spectrum (Extended Data Figs. 2, 3). Ba produces strong absorption, as do the lanthanide elements, but only in the optical region at wavelengths shorter than about 650 nm. The spectral features that we observe can therefore only be due to Sr, an element produced near the first r-process peak.

The 810-nm feature was previously proposed[12] to originate in absorption from Cs I and Te I. This identification can now be ruled out, because neither Cs I nor Te I produces strong lines in a plasma at this temperature (Extended Data Fig. 3). Much stronger lines would be expected from the ions of other elements that are co-produced with Cs (atomic number $Z = 55$) (for example, the lines from La II; see Methods).

The most abundant r-process elements are those in the first peak (Fig. 2)—elements with mass numbers ($A$) of around 80—and of these, it is Sr, Y ($Z = 39$) and Zr ($Z = 40$) that are easily detected in a low-density, roughly 4,000-K thermal plasma, because these elements have low excitation potentials for their singly charged ions. Seen in this context, the detection of Sr in AT2017gfo is not surprising, despite prior expectations that the spectra would be dominated by heavier elements[14,22]. Furthermore, the atomic levels in Sr that give the absorption lines observed at 810 nm are metastable. Photo-excitation can increase the population in these states, strengthening the 810-nm feature markedly[23] compared with the resonance blue/near-ultraviolet absorption lines. Ba and the lanthanide series contribute substantially to the total opacity of r-process material in the optical region of the spectrum (Fig. 2), yet we do not detect strong optical features. We cannot on this basis, however, easily exclude the presence of elements with mass numbers of more than 140 or so. Even if we could exclude the presence of heavier elements in the outer layers of the thermal, expanding cloud, there is no way from these early spectra of excluding the possibility that such elements could exist at lower depths or in an obscured component.

Given that a simple r-process abundance LTE model can account well for the first-epoch spectrum, we expand it to the subsequent three epochs, while the kilonova is still at least partially blackbody like. With a freely expanding explosion we expect to begin observing P Cygni lines once the outer absorbing 'atmosphere' begins to become more optically thin and attain a substantial physical radius with respect to the photospheric radius. We fit the first four epochs as a blackbody with P Cygni lines from Sr. We fit only the strongest lines in order to reduce our computational time to a manageable level, as these lines provide most of the opacity at these wavelengths. These fits are shown in Fig. 4 and offer a compelling reproduction of the spectra at all three epochs. The P Cygni model has free parameters for the velocities of the photosphere and atmosphere, which change the shape of the profile. The fit is remarkable given its simplicity and our lack of knowledge of the system geometry. We note that P Cygni emission components are always centred close to the rest wavelength of the spectral lines, so the observed wavelength of the emission line is not a free parameter. The most prominent emission component observed throughout the spectral series is centred close to 1,050 nm, and the weighted restframe centre of the near-infrared lines from Sr is also 1,050 nm. This adds to our confidence in the line identification based on the simple thermal r-process absorption model.

We further confirm our results using TARDIS ('temperature and radiative diffusion in supernovae'), extending this code's atomic database to include elements up to $_{92}$U by using the latest Kurucz line list[24] with its 2.31 million lines. Our TARDIS models produce results very similar to

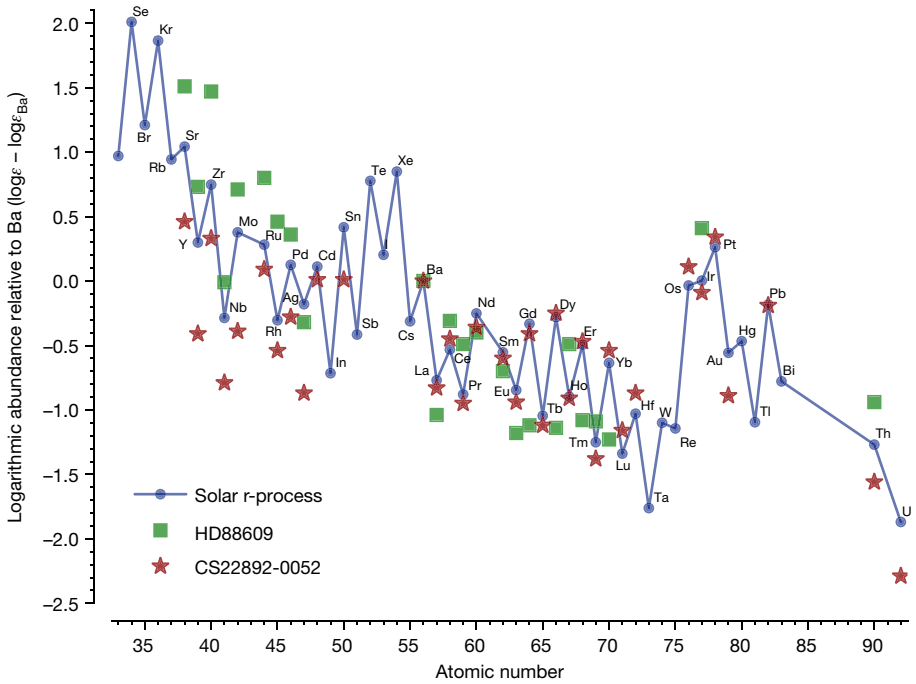



**Fig. 2 | Abundances of elements produced by the r-process.** Relative r-process abundances ($\varepsilon$) normalized to the Ba abundance are shown for the Sun and for two metal-poor stars—one, CS 22892–052, rich in heavy r-process elements[31,32], and the other, HD 88609, rich in light r-process elements[20]. These are the abundances of the elements used in the inset of Fig. 3.

our static-code models, reproducing the spectra well (Extended Data Fig. 6). In particular, the P Cygni emission/absorption structure is well reproduced as expected, confirming our LTE and MOOG modelling, and showing Sr dominating the features around 1 μm.

Given our detection of Sr, it is clearly important to consider lighter r-process elements in addition to the lanthanide elements in shaping the kilonova emission spectrum. Observations of abundances in stars in dwarf galaxies[6] suggest that large amounts of Sr are produced together

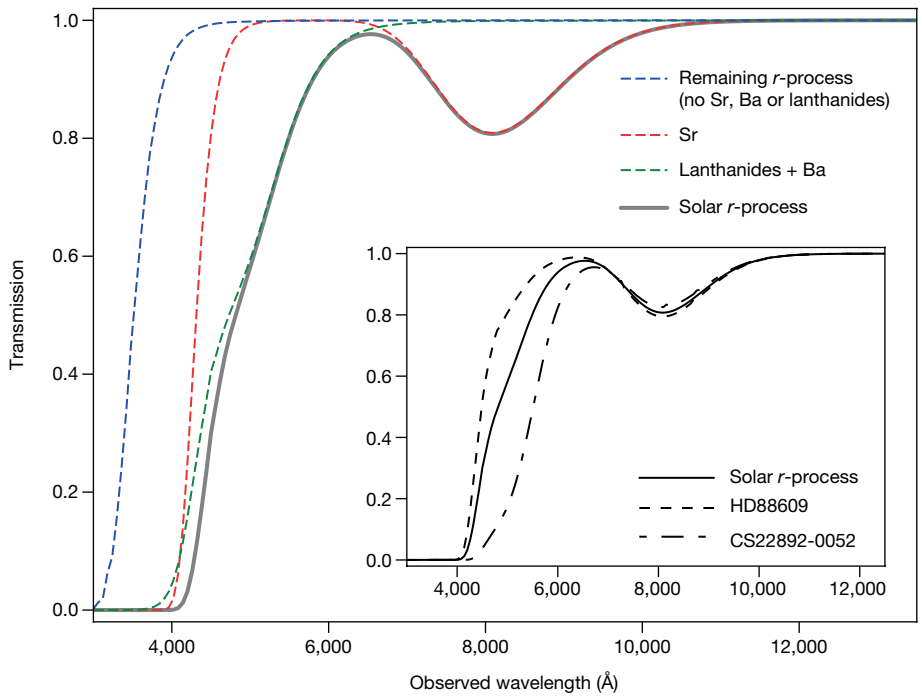



**Fig. 3 | Thermal r-process-element transmission spectrum.** These spectra are based on the lines formed in a gas in local thermal equilibrium with a temperature of 3,700 K and an electron density of $10^7$ cm$^{-3}$, broadened by 0.2$c$ and blueshifted by 0.23$c$. The spectrum produced by a solar r-process abundance ratio is plotted as a solid line. Contributions due to Sr (red dashed line), Ba and the lanthanides (green dashed line) and the remaining r-process elements (blue dashed line) are also shown. Inset, spectra resulting from a solar r-process abundance ratio (solid line), and from the abundance ratios of the metal-poor stars HD 88609 (dashed line)[20] and CS 22892–052 (dash-dotted line)[31,32].

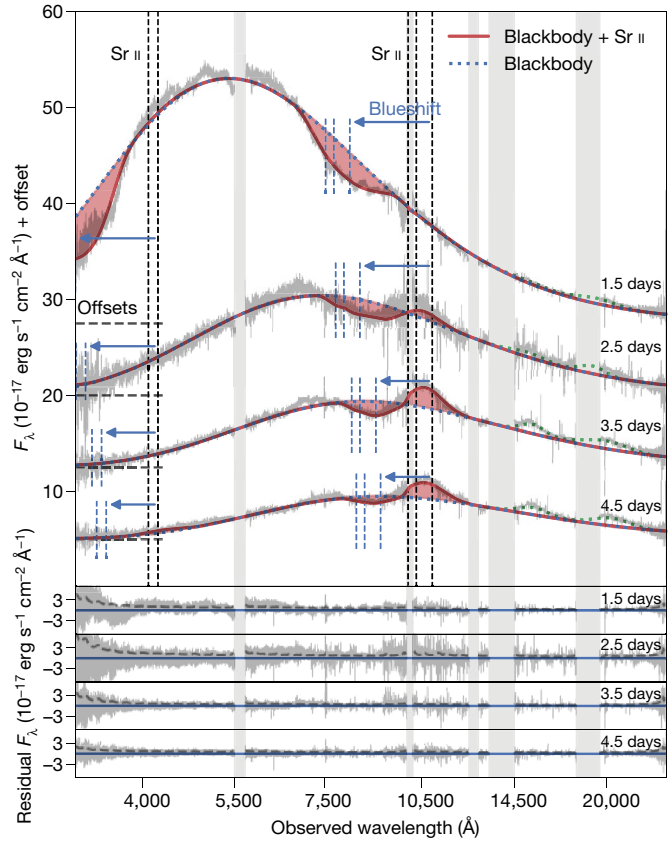



**Fig. 4 | Spectral series of AT2017gfo 1.5–4.5 days after the merger.** Data are shown in grey and have been smoothed slightly. Top panel, a model (solid red lines) consisting of a blackbody (blue dotted lines) with P Cygni profiles (red transparent fill) for the Sr lines. The rest (vertical black dashed lines) and observed (vertical blue dashed lines) positions of the model's Sr lines are shown, with the blueshift indicated by arrows. Green dotted lines show the Gaussian emission profiles added to ensure the overall continuum is not biased. A vertical offset has been applied to each spectrum for clarity, with zero flux indicated by the dashed horizontal line segments. Bottom panels show the residuals between model and data.

with Ba ($Z$ = 56) in infrequent events, implying the existence of a site that produces both light and heavy r-process elements together in quantity, as found in some models[25,26]. This is consistent with our spectral analysis of AT2017gfo and analyses of its lightcurve[27,28]. Together with the differences observed in the relative abundances of r-process Ba and Sr in stellar spectra[29], this suggests that the relative efficiencies of light and heavy r-process production could vary substantially from merger to merger.

Extreme-density stars composed of neutrons were proposed shortly after the discovery of the neutron[13], and identified with pulsars three decades later[30]. However, no spectroscopic confirmation of the composition of neutron stars has ever been made. The identification here of an element that could only have been synthesized so quickly under an extreme neutron flux provides the first direct spectroscopic evidence that neutron stars comprise neutron-rich matter.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1676-3.


1. Burbidge, E. M., Burbidge, G. R., Fowler, W. A. & Hoyle, F. Synthesis of the elements in stars. *Rev. Mod. Phys.* **29**, 547–650 (1957).
2. Siegel, D. M., Barnes, J. & Metzger, B. D. Collapsars as a major source of r-process elements. *Nature* **569**, 241–244 (2019).
3. Lattimer, J. M., Mackie, F., Ravenhall, D. G. & Schramm, D. N. The decompression of cold neutron star matter. *Astrophys. J.* **213**, 225–233 (1977).
4. Eichler, D., Livio, M., Piran, T. & Schramm, D. N. Nucleosynthesis, neutrino bursts and gamma-rays from coalescing neutron stars. *Nature* **340**, 126–128 (1989).
5. Freiburghaus, C., Rosswog, S. & Thielemann, F.-K. R-process in neutron star mergers. *Astrophys. J.* **525**, L121–L124 (1999).
6. Ji, A. P., Frebel, A., Simon, J. D. & Chiti, A. Complete element abundances of nine stars in the r-process galaxy Reticulum II. *Astrophys. J.* **830**, 93 (2016).
7. Metzger, B. D. et al. Electromagnetic counterparts of compact object mergers powered by the radioactive decay of r-process nuclei. *Mon. Not. R. Astron. Soc.* **406**, 2650–2662 (2010).
8. Barnes, J. & Kasen, D. Effect of a high opacity on the light curves of radioactively powered transients from compact object mergers. *Astrophys. J.* **775**, 18 (2013).
9. Tanvir, N. R. et al. A 'kilonova' associated with the short-duration γ-ray burst GRB 130603B. *Nature* **500**, 547–549 (2013).
10. Abbott, B. P. et al. GW170817: observation of gravitational waves from a binary neutron star inspiral. *Phys. Rev. Lett.* **119**, 161101 (2017).
11. Pian, E. et al. Spectroscopic identification of r-process nucleosynthesis in a double neutron-star merger. *Nature* **551**, 67–70 (2017).
12. Smartt, S. J. et al. A kilonova as the electromagnetic counterpart to a gravitational-wave source. *Nature* **551**, 75–79 (2017).
13. Baade, W. & Zwicky, F. Cosmic rays from supernovae. *Proc. Natl Acad. Sci. USA* **20**, 259–263 (1934).
14. Tanaka, M. & Hotokezaka, K. Radiative transfer simulations of neutron star merger ejecta. *Astrophys. J.* **775**, 113 (2013).
15. Kasen, D., Metzger, B., Barnes, J., Quataert, E. & Ramirez-Ruiz, E. Origin of the heavy elements in binary neutron-star mergers from a gravitational-wave event. *Nature* **551**, 80–84 (2017).
16. Sneden, C., Bean, J., Ivans, I., Lucatello, S. & Sobeck, J. MOOG: LTE line analysis and spectrum synthesis. *Astrophysics Source Code Library* https://www.as.utexas.edu/~chris/moog.html (2012).
17. Kerzendorf, W. E. & Sim, S. A. A spectral synthesis code for rapid modelling of supernovae. *Mon. Not. R. Astron. Soc.* **440**, 387–404 (2014).
18. Lodders, K., Palme, H. & Gail, H.-P. in *Solar System: Landolt Börnstein Group VI Astronomy and Astrophysics* Vol. 4B (ed. Trümper, J. E.) 712 (Springer, 2009).
19. Bisterzo, S., Travaglio, C., Gallino, R., Wiescher, M. & Käppeler, F. Galactic chemical evolution and solar s-process abundances: dependence on the $^{13}$C-pocket structure. *Astrophys. J.* **787**, 10 (2014).
20. Honda, S., Aoki, W., Ishimaru, Y. & Wanajo, S. Neutron-capture elements in the very metal-poor star HD 88609: another star with excesses of light neutron-capture elements. *Astrophys. J.* **666**, 1189–1197 (2007).
21. Sneden, C. et al. Evidence of multiple r-process sites in the early galaxy: new observations of CS 22892–052. *Astrophys. J.* **533**, L139–L142 (2000).
22. Kasen, D., Badnell, N. R. & Barnes, J. Opacities and spectra of the r-process ejecta from neutron star mergers. *Astrophys. J.* **774**, 25 (2013).
23. Jeffery, D. J. & Branch, D. in *Supernovae, Jerusalem Winter School for Theoretical Physics* Vol. 6 (eds Wheeler, J. C., Piran, T. & Weinberg, S.) 149 (World Scientific, 1990).
24. Kurucz, R. L. Including all the lines: data releases for spectra and opacities. *Can. J. Phys.* **95**, 825–827 (2017).
25. Wanajo, S. et al. Production of all the r-process nuclides in the dynamical ejecta of neutron star mergers. *Astrophys. J.* **789**, L39 (2014).
26. Just, O., Bauswein, A., Pulpillo, R. A., Goriely, S. & Janka, H.-T. Comprehensive nucleosynthesis analysis for ejecta of compact binary mergers. *Mon. Not. R. Astron. Soc.* **448**, 541–567 (2015).
27. Drout, M. R. et al. Light curves of the neutron star merger GW170817/SSS17a: implications for r-process nucleosynthesis. *Science* **358**, 1570–1574 (2017).
28. Tanvir, N. R. et al. The emergence of a lanthanide-rich kilonova following the merger of two neutron stars. *Astrophys. J.* **848**, L27 (2017).
29. Hansen, C. J., Montes, F. & Arcones, A. How many nucleosynthesis processes exist at low metallicity? *Astrophys. J.* **797**, 123 (2014).
30. Hewish, A., Bell, S. J., Pilkington, J. D. H., Scott, P. F. & Collins, R. A. Observation of a rapidly pulsating radio source. *Nature* **217**, 709–713 (1968).
31. Sneden, C., Lawler, J. E., Cowan, J. J., Ivans, I. I. & Den Hartog, E. A. New rare earth element abundance distributions for the Sun and five r-process-rich very metal-poor stars. *Astrophys. J.* **182** (Suppl.), 80–96 (2009).
32. Sneden, C. et al. The extremely metal-poor, neutron capture-rich star CS 22892-052: a comprehensive abundance analysis. *Astrophys. J.* **591**, 936–953 (2003).

# Methods

### Spectral synthesis

We used different codes to compute synthetic absorption spectra, namely MOOG[16,33] v. 2014 and our own single-temperature and single-density LTE code. In addition, we verified our results using the TARDIS supernova spectral synthesis code. For the first two codes, we used line lists gathered from the literature (see Supplementary Information). For the TARDIS modelling, we used the line lists of Kurucz[24]. Our codes yield consistent results with the different line lists.

MOOG is a synthetic spectrum code normally used to generate synthetic absorption spectra of photospheres in cool stars under the assumption of local thermodynamic equilibrium. It requires a model atmosphere that dictates how temperature, gas pressure and electron density behave in different layers of the surface gas. Here we adopt Kurucz model atmospheres[34]. The second requirement is a line list that contains the rest wavelength of the absorption transition, the element or ion in which the transition takes place, the excitation potential of the lower level, and the oscillator strength. The atomic data are based on refs. [31,35–40] with updates from the National Institute of Standards and Technology (NIST). The strengths of the absorption features are calculated solving radiative transfer equations with a plane parallel treatment of the atmospheres, assuming that the velocity distribution is Maxwellian, and that excitations and ionizations are described by the Boltzmann and Saha equations, respectively. The line/wing damping follows a scaled Unsöld approximation and the source function follows a simple blackbody, while scattering (on H, He and $e^-$) enters mainly through opacity terms.

Our own code assumes only a gas in LTE without scattering, and that the Boltzmann and Saha equations can be used to obtain the ionization and excitation state of each element individually. We then use the line lists above and level information from NIST to determine the relative strengths of the lines. We adopt a fiducial electron density of $\log n_e = 7.8$, based on the mean density of $0.04 M_\odot$ (where $M_\odot$ is the mass of the Sun) of singly ionized material in a sphere with the area of the best-fit blackbody. The density of the atmosphere is almost certainly lower than this.

Demonstrating that the MOOG models and our LTE calculations are reasonably comparable, for the MOOG models an effective temperature ($T_{\rm eff}$) at the surface of the photosphere of roughly 5,500 K and a surface gravity of log = 0 (following the temperature and density profiles in the Kurucz model atmospheres) give rise to a temperature of 3,800 K and an electron density of $n_e = 10^7$ cm$^{-3}$ within the photosphere. Absorption lines from lanthanide ions are believed to be an important source of opacity owing to transitions with unknown oscillator strengths. For an LTE plasma, it is likely that such lines are important and create a complex continuum[15,22]. However, the lanthanide opacity is extremely high in the ultraviolet and blue regions of the spectrum. The fact that we detect blue emission in the spectrum of AT2017gfo is already a strong indication that lanthanide elements do not dominate the early-continuum spectrum, as suggested previously[41,42]. Furthermore, the infrared feature arises from levels that may be overpopulated owing to optical pumping, enhancing the strength of this feature further with respect to the line-generated continuum at these wavelengths.

Synthetic spectra are generated using both codes on the basis of line lists containing r-process elements capable of producing strong features in an LTE plasma at these temperatures. We include all elements from $_{33}$As up to $_{83}$Bi, as well as $_{90}$Th and $_{92}$U. We do not include the elements $_{35}$Se, $_{36}$Br, $_{37}$Kr, $_{53}$I and $_{54}$Xe as they produce no strong or moderate lines at these temperatures and are rarely detected in stellar spectra[43]; these elements have first excitation energies above 5.97 eV for their neutral and singly charged ions, giving a fractional population less than $10^{-8}$ at our fiducial temperature. Neither do we include elements with no stable isotopes ($_{43}$Tc and $_{61}$Pm), or any molecules. The absorption-line profiles are dominated by the velocity and density distribution of the expanding atmosphere.

Our line lists contain the strongest lines for LTE spectra at these temperatures. Because we are interested in finding strong, isolated lines, this procedure should effectively capture all lines that could realistically be candidates.

### Could large numbers of weak lines dominate the opacity?

The opacity of the kilonova is dominated by absorption lines. The list of lines that we use for MOOG (see references above) has most of the strong lines in common with the Kurucz list[44] that we use for the TARDIS modelling. The results we retrieve from the different techniques and line lists are a useful check on the robustness of the modelling methodologies. Both methods yield consistent results, indicating that the overall result presented here is robust to the selection of the specific line list and the modelling method chosen. We note that a feature at about 810 nm is also produced in the spectral synthesis analysis of ref. [14], where lists comprising known lines are also used. This feature (M. Tanaka, private communication) is produced primarily by the same Sr II lines we identify in this work.

The major caveat in identifying line features is the possibility that missing lines could have a larger influence on the broad spectral shape than the predicted effect from known lines. Of particular concern are the large numbers of unknown lines from the lanthanide elements that are likely to dominate the line-expansion opacity[22,45]. Although we argue here that our line lists are reasonably complete in strong lines at these temperatures and densities (and given that they are used for modelling stars with similar temperatures and densities, this makes sense), it is possible that a very large number of weaker lines could contribute.

However, the line-forming region of the kilonova is likely to be physically extended, covering a substantial fraction of the kilonova radius, particularly in the near-infrared. The presence of a P Cygni profile at around 1 μm supports the idea that a substantial volume (though not mass) of the kilonova must be largely optically thin at this wavelength. The mass absorption coefficient of the Sr II lines at around 1.05 μm peaks at about $4 \times 10^3$ cm$^2$ g$^{-1}$ for lines with a full width at half maximum (FWHM) of $0.01c$, a temperature of 5,000 K and a density of $10^{-13}$ g cm$^{-3}$. This is at least two orders of magnitude higher than the mean value obtained for lanthanides such as Ce and Nd in the optically thin limit using the Kurucz line lists. Given that the line lists for these elements are likely to be highly incomplete at these wavelengths, we extrapolate the value of the Ce line opacity of the Vienna Atomic Line Database (VALD) lines at 9,000 Å to be roughly 1.05 μm, which should give a similar opacity to the line lists calculated in ref. [22] with the autostructure code. When the lines are extremely optically thick, within the bulk of the kilonova in the first days, the Ce opacity is about 10 cm$^2$ g$^{-1}$ (compare with ref. [22]). In the optically thin regime in the outer layers, the Ce line opacity rises by about two orders of magnitude. Using this optically thin extrapolation of the Ce lines, the Sr II opacity is still a factor of four to five times higher, not including abundance effects that are likely to make the Sr line stronger still. We show an example of this effect by calculating the expansion opacity for a low-optical-depth plasma in Extended Data Fig. 5. That calculation is purely illustrative, showing how the Sr lines can dominate the opacity when the gas has low optical depth. For a self-consistent model calculation, see the TARDIS model spectra in Extended Data Fig. 6.

### Spectral modelling

In the spectra we identify what appear to be two separate emission components: first, a nearly blackbody spectrum modified by absorption features that appears to cool over time; and second, an emission component at redder wavelengths that increases in strength relative to the first component with time. These two components do not necessarily arise because of discrete ejection mechanisms, but may reflect the fact that different parts of the spectrum probe different physical depths and thus physical conditions, through the wavelength-dependent expansion opacity[8,46]. Here we focus only on the thermal component in the blue part

of the spectrum and model it as a blackbody with an extended envelope. We model the second component with Gaussian emission lines in order not to bias the overall continuum fit at shorter wavelengths, but do not interpret them. However, these features clearly provide important information on the composition of the plasma and must be addressed in future studies.

The expansion velocity of the gas can be inferred from the expansion of the blackbody from the time of the explosion. Owing to the optical thickness of a blackbody, we would only be presented with the front face of the explosion. Consequently, pure absorption features in the spectrum should be blueshifted by the mean Doppler shift induced by the expansion speed of the gas. Conservatively, we allow 0.1–0.3$c$ as the range of the blueshift[11,47,48], a value that depends on the details of the geometry of the system, and thus we restrict our search for lines in the first epoch to rest wavelengths of 350 nm and 810 nm multiplied by 1.1–1.3.

At the densities of the ejecta, the dominant source of opacity is expansion opacity[15,22]. This effect is able to establish an apparent thermalization through photo-equilibration of the states[49]. With wavelength-dependent opacity, the physical depth traced at each wavelength varies. Because the large majority of lines are at the blue end of the spectrum, the expansion opacity there will be higher and, conversely, the physical depth shallower. This causes the relative strength of ultraviolet/near-infrared lines to change compared with the pure LTE transmission values, with bluer absorption lines being less prominent relative to near-infrared ones. Additionally, because the population of states is photo-equilibrated, metastable states will be enhanced relative to non-metastable, as compared with LTE values[23]. It is therefore impossible, primarily because of the strongly wavelength-dependent opacity, to use a simple comparison of LTE line strengths across very different wavelengths. Instead, we use independent optical depth parameters ($\tau$) for the two absorption feature fits here. We also use the TARDIS code (see below) to achieve a more self-consistent treatment with moving atmosphere, line-expansion opacity, which shows the simultaneous presence of the Sr II features at around 0.4 μm and 1 μm.

### P Cygni modelling

The expansion velocity of the photosphere is very high (0.2–0.3$c$). At the measured temperature of the photosphere, the thermal widths of individual lines are very narrow compared with the gross velocity structure. This means that the resonance region is very small and the Sobolev approximation can be used in the Elementary Supernova model as a prescription for the absorption structure near isolated lines[23]. We use the implementation of the P Cygni profile in the Elementary Supernova from https://github.com/unoebauer/public-astro-tools, where the profile is parametrized in terms of the rest wavelength, $\lambda_0$, the optical depth of the line, $\tau$, two scaling velocities for the radial dependence of $\tau$, the photospheric velocity, and the maximal velocity of the ejecta. The latter two parameters specify the velocity stratification. The expansion velocity of the photosphere is simultaneously used for the relativistic Doppler correction to the blackbody temperature. In addition, because the implementation of the P Cygni profile that we are using does not include the relative population of the states in the transition, we have included a parameter for enhancement/suppression of the P Cygni emission component.

For practical reasons, we cannot fit all lines simultaneously. However, fortunately, a handful of lines provides most of the opacity. Because the relative opacity dictates the apparent strengths of the lines, we divide the spectrum into ultraviolet/blue and red/infrared regions to find the lines that will be strongest in their respective spectral region. We do this because the opacity changes so severely from the infrared to the optical (Fig. 3). We make the division at 600 nm where the opacity increases sharply; however, choosing 550 nm or 700 nm makes no difference. We then include the strongest lines in each region (all lines with a minimum strength of 20% of the strongest line). The resulting lines are the strong resonance lines from the ground state of Sr II at 407.771 nm and 421.552 nm, and the lines from the Sr II $4p^6 4d$ metastable states at 1,032.731 nm, 1,091.489 nm and 1,003.665 nm. These lines are all modelled using the same P Cygni profile prescription, where the relative strengths of each of the lines in the two absorption complexes are set by the LTE relations, and despite the relative simplicity of the analysis, this approach provides a surprisingly good fit to the data.

The final model that we use to fit the spectrum is a relativistically corrected blackbody photosphere absorbed by an expanding atmosphere, containing the five above-mentioned Sr II transitions, described by independent optical depths for the infrared and ultraviolet lines. The ratios of the lines internally in each set are defined by their LTE strengths. In the fitting model we also use two additional Gaussian emission lines at long wavelengths from the second emission component in order not to bias the long-wavelength continuum fit. The best-fit parameters and their associated errors are found by sampling the posterior probability distributions of the parameters, assuming flat priors on all parameters. The fitting framework used is LMFIT[50] and the sampling is done using emcee[51]. We initiate 100 samplers, each sampling for 1,000 steps. We discard the first 100 steps as a burn-in phase of the Markov chain Monte Carlo (MCMC) chains. We use the median of the marginalized posterior probability distribution as the best-fit values, and the 16th and 84th percentiles as the uncertainties. The best-fit models are shown in Fig. 4. The objective function, being highly nonlinear, causes the posterior probability distributions to be highly complex and the best-fit values difficult to optimize. However, the peaks of the distributions are well centred, meaning that the best-fit values are well constrained, regardless of the complexity of the posterior probability distribution.

### Expansion velocity evolution

The fits constrain two independent parameters that can be used to infer the velocity of the ejected material: the photospheric expansion velocity, used to determine the width of the P Cygni line profile; and the blackbody radius, which scales with the square root of the observed luminosity and can be converted to an expansion velocity on the basis of the time of observation. These two parameters are uncorrelated, as supported by the MCMC posterior probability function samples, and therefore constitute two independent measurements of the same physical quantity. We show a plot of the evolution of these two parameters in Extended Data Fig. 4. The correspondence between the two estimates of the expansion velocity is striking, especially given that the ratio of the estimates is geometry dependent, and we have assumed only simple spherical symmetry here. Only the first epoch shows a somewhat discrepant value, and there we do not expect a P Cygni model to be entirely applicable. This close correspondence between the two independent measures and the reasonable values inferred further supports the validity of the line identification and the overall model.

### TARDIS modelling

TARDIS[17] is a Monte Carlo radiative-transfer spectral synthesis code, in which photons are essentially propagated through an expanding atmosphere. Each photon will at any point have a probability of being absorbed by an atomic transition, this probability being based on the wavelength of the photon, the strength of the line, and the density of atomic species and electron populations. A synthetic spectrum can then be constructed by collecting the emergent photons.

To generate the synthetic spectra using TARDIS, we set up the physical models using the inferred photospheric expansion velocities at the observed epochs. For homologously expanding ejecta, the velocities of the atmosphere layers are at all times specified by the outer-edge expansion and the photospheric expansion. We use the measured photospheric expansion velocity as the inner expansion velocity and select the outer atmospheric velocity such that the bluest edge of the developed absorption profiles in the synthetic spectra match the observed ones. At present, TARDIS supports only spherically symmetric explosions, so

for simplicity we adopt this geometry. The kilonova ejecta are in most cases likely to be asymmetric, owing to the preferential motion of the mass in the plane of the orbit of the two neutron stars. The neglect of deviation from spherical symmetry most likely affects the absorption profiles and the inferred mass in the atmosphere, as we could potentially only be seeing ejecta in a cone. Additionally, TARDIS assumes a single photospheric velocity across the entire wavelength range. Owing to the strong wavelength dependence of the opacity, as discussed earlier, the depth at which the photons escapes varies across the spectral coverage. Therefore, the same reservations about inferring the mass in a given shell at a given wavelength applies to the TARDIS simulations. This can be seen in effect when choosing an ejecta density that matches the absorption feature at 350 nm, because then the strength of the 810-nm absorption feature is greatly overpredicted. Conversely, selecting an ejecta density that matches the 810-nm absorption feature underpredicts the strength of the 350-nm absorption.

At each epoch, the temperature of the photosphere is chosen so that an atmosphere with no lines returns a blackbody-like spectrum that is similar to the best-fit blackbody found in simple P Cygni model fits. Both the excitation and the ionization structure of the elements in the atmosphere are set according to LTE, where we assume for simplicity a constant temperature throughout the atmosphere. This approach does not capture optical pumping of metastable states and other non-LTE effects that will change the population of the upper levels.

For the input abundances, we use the solar *r*-process abundance ratio as shown in Fig. 2, starting from $_{31}$Ga. We run the simulation in three steps, consecutively including heavier elements. For the first set of simulations, we include only the elements from $_{31}$Ga to $_{37}$Rb and, as can be seen in Fig. 2, no lines cause a substantial deviation from a pure blackbody. Next we include $_{38}$Sr, which forms the strong feature observed centred at 810 nm in the first epoch, almost exclusively owing to the three strong Sr II lines at around 1 μm. Finally we run the same simulation, including all elements from $_{31}$Ga to $_{92}$U. The feature at 810 nm is unaffected by the inclusion of the heavier elements.

For the density, we initially adopt a power-law density structure of the ejecta, parametrized in terms of velocity and epoch: $\rho(v,t) = \rho_0(t_0)^3(v/v_0)^n$. We find that the line shapes depend on the assumed slope, where for steeper slopes a larger fraction of the line absorption is closer to the line centre. We specify a density profile of $n = -3$, as in ref. [52], as this supported by the theoretical models and seems to reproduce the absorption profiles relatively well. As also investigated in ref. [15], there is some freedom in the choice of slope, as it is not well constrained from a modelling perspective and could have different values depending on the matter-ejection mechanism.

Adopting a single $\rho_0$ across all four epochs, with $n = -3$, does not yield synthetic spectra that match the observed spectra well around the 810-nm $_{38}$Sr absorption feature across the epochs. If $\rho_0$ is chosen to reproduce the strength of the $_{38}$Sr absorption feature of the first epoch, the strength of the absorption feature is greatly overpredicted in the later epochs using the same composition and assuming homology; the ejecta density has to be scaled down by a factor of five in the subsequent epochs to match the spectrum. In other words, the observed mass of Sr in the optically thin part of the spectrum inferred from the TARDIS model for the first-epoch spectrum appears to be much larger than for the later epochs. Specifically, atmosphere masses of $5 \times 10^{-5} M_\odot$, $1 \times 10^{-5} M_\odot$, $1.2 \times 10^{-5} M_\odot$ and $1.3 \times 10^{-5}$ of $_{38}$Sr are required to reproduce the observed absorption feature at 810 nm for the first four epochs respectively.

These numbers should be treated with some caution as these are derived masses assuming spherical symmetry, a fixed photospheric velocity, and no correction for light travel time effects. They must be interpreted as lower limits to the total amount of material ejected, as they trace only the matter between the photospheric front and the outer atmosphere. Using the assumed solar abundances, these masses correspond to this atmosphere having approximately 1% of the total ejecta mass inferred from lightcurve modelling[15].

The TARDIS models also constrain the amount of the heavier r-process elements present in the outer, transparent layers of the ejecta. Using the solar r-process abundances with the inclusion of the heaviest elements, the TARDIS synthetic spectra exhibit almost continuous absorption up to around 6,000 Å, which is not seen in the observed spectra. This point was also touched upon earlier. The exact limit to the amount of heavy r-process material in the outer layers is difficult to infer accurately, on the basis of the simple models used, but our modelling indicates that the ratio of heavy to light element abundance in this layer is much smaller than the solar r-process ratio. This conclusion is consistent with the inference made by other authors on the basis of the early blue colour of the continuum spectrum[41,42].

The inability of a single composition and density to reproduce the spectra across the first four epochs may hint at a change in the elemental abundance ratios as the photosphere recedes further into the ejecta.

The TARDIS models demonstrate that an isolated feature observed at 810 nm can be produced by Sr and that no other known lines form this feature. Additionally, the models hint at a possible variation in the abundances as the deeper layers of the ejecta component are exposed, in line with what is suggested by some models of neutron-star mergers[53].

### Exclusion of Cs I and Te I identification

The Cs I 6*s*–6*p* resonance transitions[12] would of course require Cs I to be present in the gas. But because Cs has the lowest first ionization potential of any element, the singly charged ions of other elements inevitably synthesized with Cs, such as La II, Eu II and Gd II, are millions of times more abundant than Cs I in an LTE plasma at close to the observed blackbody temperatures. This problem is even worse at temperatures that produce substantial strong lines from Te I. These other elements will cause absorption lines that are at least two orders of magnitude stronger in the same wavelength region as the proposed Cs and Te lines—for example, the 706.62 nm, 742.66 nm or 929.05 nm lines of La II, Eu II and Gd II respectively, to name one of each. The same argument holds for the excited-state transition of Te I, which has a very high excitation energy of 5.49 eV; the relative population of the Te I excited state is extremely low, less than $10^{-7}$. Thus, no realistic scenario exists in which lines from either of these species can be detected without being dominated by lines from other elements that are orders of magnitude stronger.

## Data availability

Work in this paper was based on observations made with European Space Observatory (ESO) telescopes at the Paranal Observatory under programmes 099.D-0382 (principal investigator E. Pian), 099.D-0622 (principal investigator P. D'Avanzo), 099.D-0376 (principal investigator S. J. Smartt) and 099.D-0191 (principal investigator A. Grado). The data are available at http://archive.eso.org.


33. MOOG spectral synthesis code. https://www.as.utexas.edu/~chris/moog.html (C. Sneden, 2017).
34. Castelli, F. & Kurucz, R. L. New grids of ATLAS9 model atmospheres. https://arxiv.org/abs/astro-ph/0405087 (2004).
35. Biémont, E. & Quinet, P. Recent advances in the study of lanthanide atoms and ions. *Physica Scripta* **T105**, 38 (2003).
36. Den Hartog, E. A., Lawler, J. E., Sneden, C. & Cowan, J. J. Improved laboratory transition probabilities for Nd II and application to the neodymium abundances of the Sun and three metal-poor stars. *Astrophys. J.* **148** (Suppl.), 543–566 (2003).
37. Lawler, J. E., Bonvallet, G. & Sneden, C. Experimental radiative lifetimes, branching fractions, and oscillator strengths for La II and a new determination of the solar lanthanum abundance. *Astrophys. J.* **556**, 452–460 (2001).
38. Lawler, J. E., Wickliffe, M. E., den Hartog, E. A. & Sneden, C. Improved laboratory transition parameters for Eu II and application to the solar europium elemental and isotopic composition. *Astrophys. J.* **563**, 1075–1088 (2001).
39. Lawler, J. E., Wickliffe, M. E., Cowley, C. R. & Sneden, C. Atomic transition probabilities in Tb II with applications to solar and stellar spectra. *Astrophys. J.* **137** (Suppl.), 341–349 (2001).
40. Lawler, J. E., Den Hartog, E. A., Sneden, C. & Cowan, J. J. Improved laboratory transition probabilities for Sm II and application to the samarium abundances of the Sun and three r-process-rich, metal-poor stars. *Astrophys. J.* **162** (Suppl.), 227–260 (2006).

41. McCully, C. et al. The rapid reddening and featureless optical spectra of the optical counterpart of GW170817, AT 2017gfo, during the first four days. *Astrophys. J.* **848**, L32 (2017).
42. Chornock, R. et al. The electromagnetic counterpart of the binary neutron star merger LIGO/Virgo GW170817. IV. Detection of near-infrared signatures of r-process nucleosynthesis with Gemini-South. *Astrophys. J.* **848**, L19 (2017).
43. Sneden, C., Cowan, J. J. & Gallino, R. in *Chemical Abundances in the Universe: Connecting First Stars to Planets* Vol. 265 (eds Cunha, K., Spite, M. & Barbuy, B.) 46–53 (IAU Symposium, 2010).
44. Kurucz line list. http://kurucz.harvard.edu/linelists/gfnew/gfall08oct17.dat.
45. Tanaka, M. et al. Properties of kilonovae from dynamical and post-merger ejecta of neutron star mergers. *Astron. Astrophys.* **852**, 109 (2018).
46. Karp, A. H., Lasher, G., Chan, K. L. & Salpeter, E. E. The opacity of expanding media—the effect of spectral lines. *Astrophys. J.* **214**, 161 (1977).
47. Shappee, B. J. et al. Early spectra of the gravitational wave source GW170817: evolution of a neutron star merger. *Science* **358**, 1574–1578 (2017).
48. Waxman, E., Ofek, E., Kushnir, D. & Gal-Yam, A. Constraints on the ejecta of the GW170817 neutron-star merger from its electromagnetic emission. *Mon. Not. R. Astron. Soc.* **481**, 3423–3441 (2018).
49. Pinto, P. A. & Eastman, R. G. The physics of type IA supernova light curves. II. Opacity and diffusion. *Astrophys. J.* **530**, 757–776 (2000).
50. Newville, M. et al. Lmfit: non-linear least-square minimization and curve-fitting for Python. *Astrophysics Source Code Library* https://zenodo.org/record/11813#.XX-EoS3MxuU (2016).
51. Foreman-Mackey, D., Hogg, D. W., Lang, D. & Goodman, J. emcee: the MCMC hammer. *Publ. Astron. Soc. Pacific* **125**, 306 (2013).
52. Tanaka, M. et al. Kilonova from post-merger ejecta as an optical and near-Infrared counterpart of GW170817. Publ. *Astron. Soc. Japan* **69**, 102 (2017).
53. Perego, A. et al. Neutrino-driven winds from neutron star merger remnants. *Mon. Not. R. Astron. Soc.* **443**, 3134–3156 (2014).

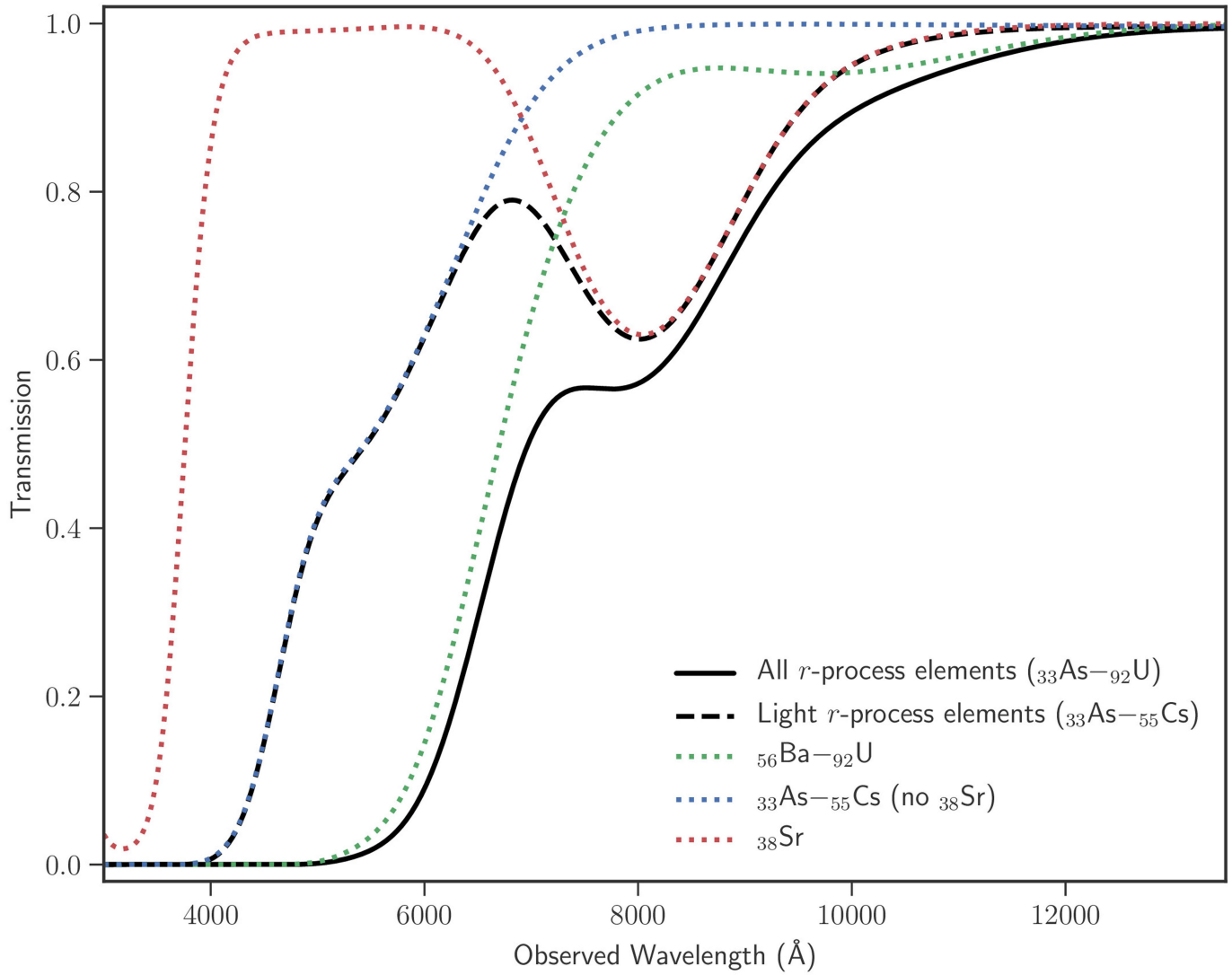


**Acknowledgements** We thank M. Tanaka for revisiting his previous analysis for us and for access to his spectra and line lists. We thank J. Hjorth and N. Rea for discussions. We thank the ESO Director General for allocating Director's Discretionary Time to this programme, and the ESO operation staff for support. D.W., D.B.M., and J.S. are supported in part by Independent Research Fund Denmark grant DFF-7014-00017. The Cosmic Dawn Center is funded by the Danish National Research Foundation under grant number 140. C.J.H. acknowledges support from the 'ChETEC' COST Action (CA16117), supported by COST (European Cooperation in Science and Technology). A.A. is supported by the European Research Council (ERC) through ERC Starting Grant 677912 EUROPIUM. A.A. and A.B. are supported by the Sonderforschungsbereich SFB 1245 'Nuclei: From Fundamental Interactions to Structure and Stars'. A.B. and A.K. are supported by the Sonderforschungsbereich SFB 881 'The Milky Way System' (subprojects A03, A05, A10 and A11) of the German Research Foundation (DFG). A.B. is supported by the ERC through ERC Starting Grant 759253 GreatMoves, and acknowledges support from the Klaus Tschira Foundation. S.C. acknowledges partial funding from Agenzia Spaziale Italiana-Istituto Nazionale di Astrofisica grant I/004/11/3. G.L. is supported by a research grant (19054) from Villum Fonden. K.E.H. acknowledges support by a Project Grant (162948-051) from The Icelandic Research Fund. A.J.L. acknowledges funding from the ERC under grant agreement 725246, and from the UK Science and Technologies Facilities Council (STFC) via grant ST/P000495/1. E.P. acknowledges funding from the Agenzia Spaziale Italiana (ASI) INAF grant I/088/06/0, and from the INAF project 'Gravitational Wave Astronomy with the first detections of aLIGO and aVIRGO experiments'.

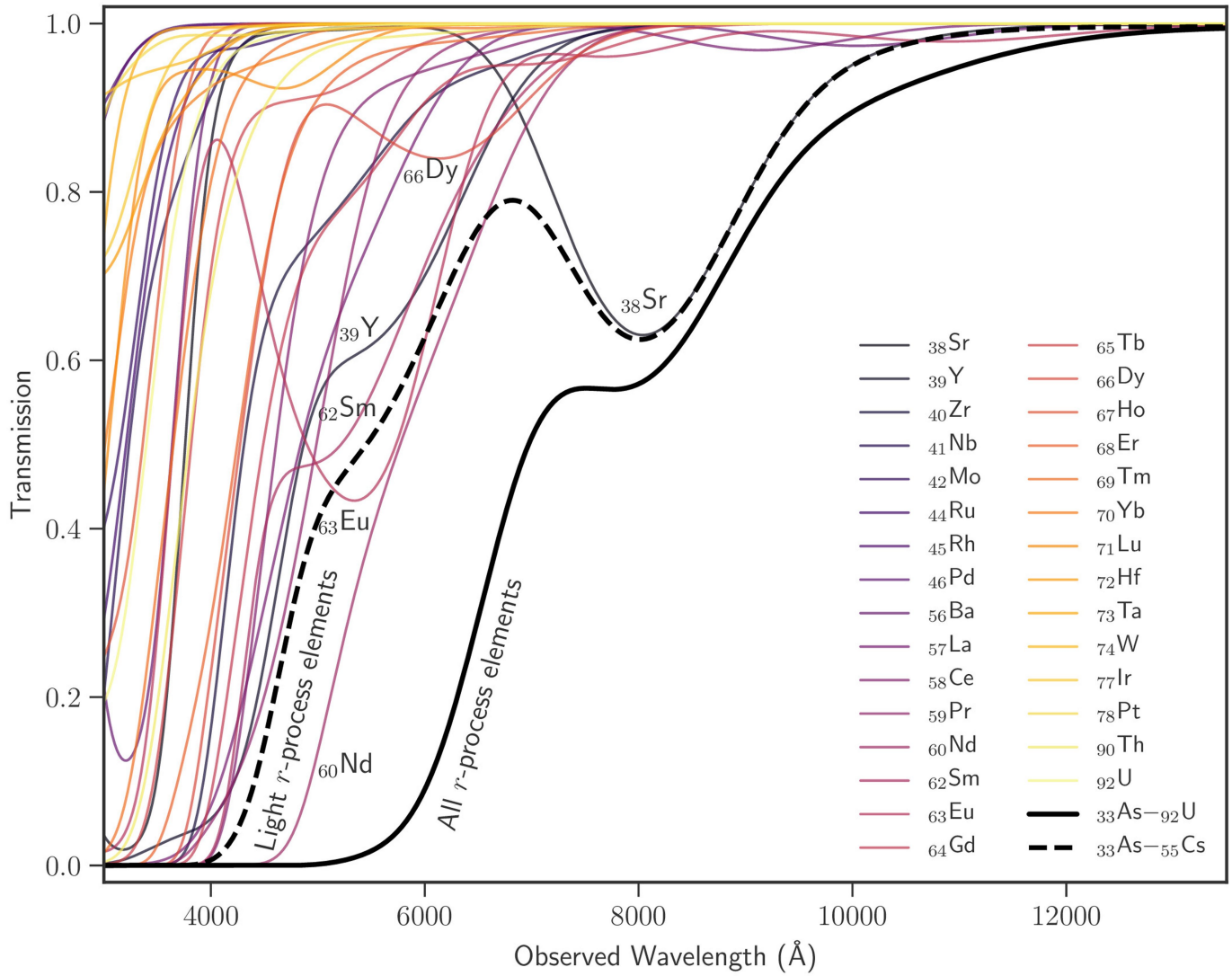

**Author contributions** D.W., C.J.H. and J.S. were the primary drivers of the project; A.K., D.B.M., J.P.U.F. and A.C.A. were involved in discussions that developed the understanding of the physical processes. All authors contributed to discussions and to editing of the paper. D.W. did the initial blackbody with absorber fits to the first-epoch spectrum, created Figs. 1, 3 and Extended Data Fig. 3, 5, made the initial line identification, recognised the P Cygni profiles in the later epochs, wrote the LTE code, and was the primary author of the main text. C.J.H. computed the initial models and synthetic spectra with MOOG, and generated the MOOG spectra for HD 88609 and CS 22892–052. C.J.H. and A.K. produced the MOOG spectra from 3,000 Å to 20,000 Å for the kilonova template photosphere for all heavy elements. C.J.H. wrote the sections on MOOG spectrum synthesis and large parts of the text on nucleosynthesis. A.K. provided the line lists. J.S. reduced and processed all the X-shooter data, produced the P Cygni fitting codes and fit the P Cygni profiles to all epochs, as well as extending the TARDIS code to include the Kurucz line lists and implementing the TARDIS modelling. J.S. also produced Fig. 3 and Extended Data Figs. 1, 2, 6, and wrote the related Methods sections and a substantial part of the main text.

**Competing interests** The authors declare no competing interests.

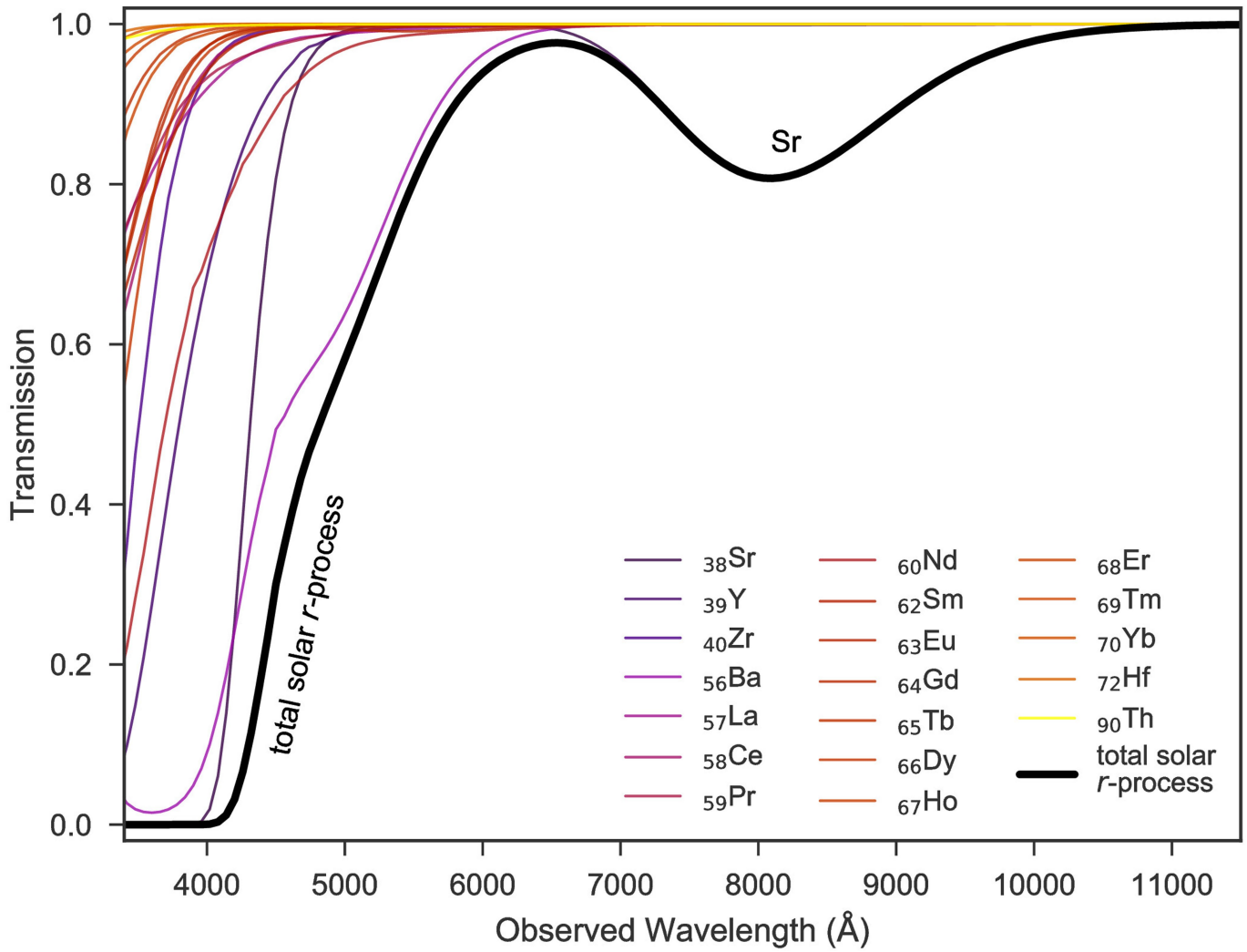
**Additional information**
**Supplementary information** is available for this paper at https://doi.org/10.1038/s41586-019-1676-3.
**Correspondence and requests for materials** should be addressed to D.W.
**Reprints and permissions information** is available at http://www.nature.com/reprints.



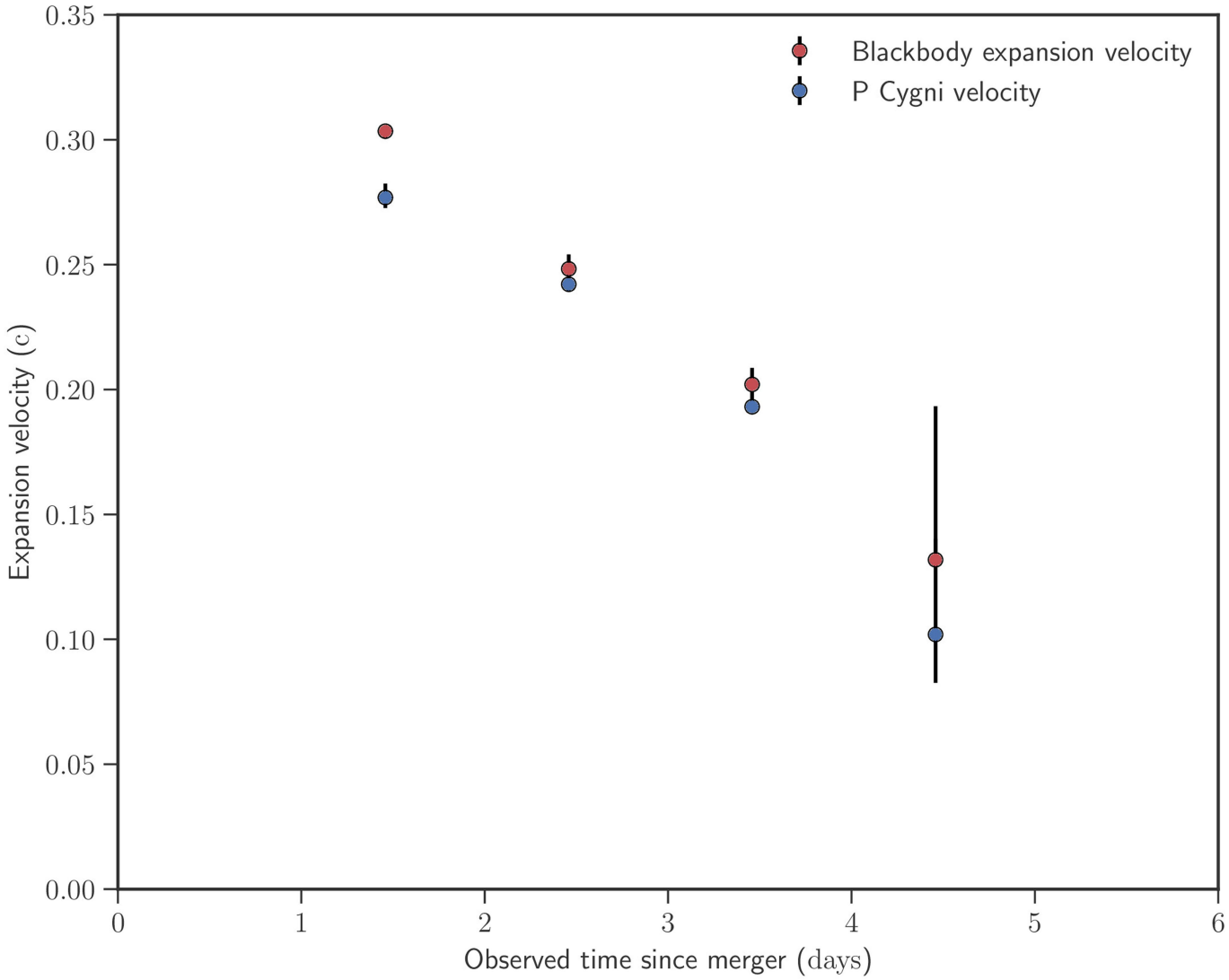

**Extended Data Fig. 1 | Synthetic r-process-element transmission spectra.** These spectra were generated using MOOG, in which the relative abundances are based on solar r-process abundances. The spectra were blueshifted, broadened and normalized as in Fig. 3. The solid black line is the total transmission spectrum for an atmosphere containing all the r-process elements ($_{33}$As to $_{92}$U). The dashed black line is the same spectrum, but including only the light r-process elements ($_{33}$As to $_{55}$Cs). The contributions from different subsets of species are also shown. The green dotted line shows the heavy r-process elements ($_{56}$Ba to $_{92}$U); the blue dotted line shows the light r-process elements ($_{33}$As–$_{55}$Cs) excluding Sr, which is shown as a red dotted line. This plot shows how Sr stands out in absorption, regardless of the composition of the material. The normalization is arbitrary and different to the LTE equivalent in Fig. 3 for display reasons.



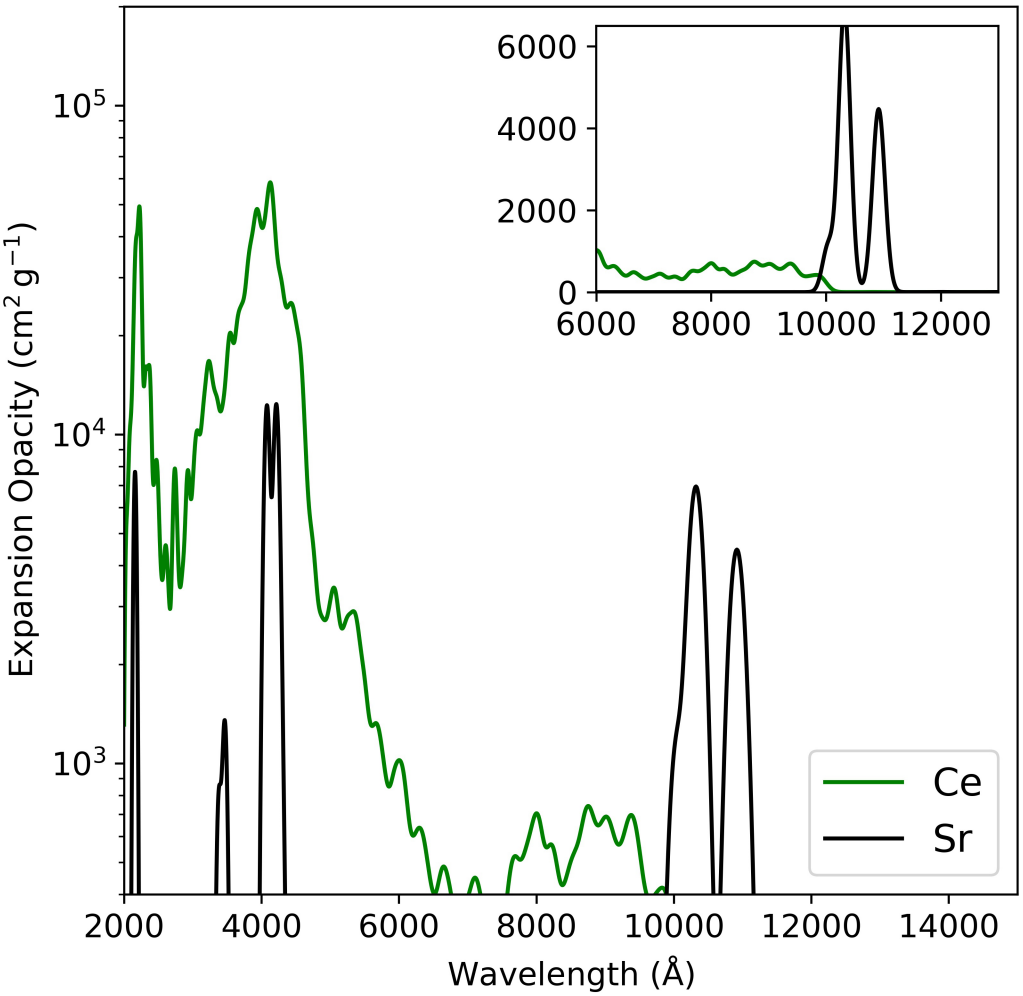

**Extended Data Fig. 2 | Synthetic r-process transmission spectra.** The spectra were generated with MOOG and are similar to those shown in Extended Data Fig. 1, except that all element contributions are displayed individually. The elements that contribute most at the reddest wavelengths are noted within the plotted line.



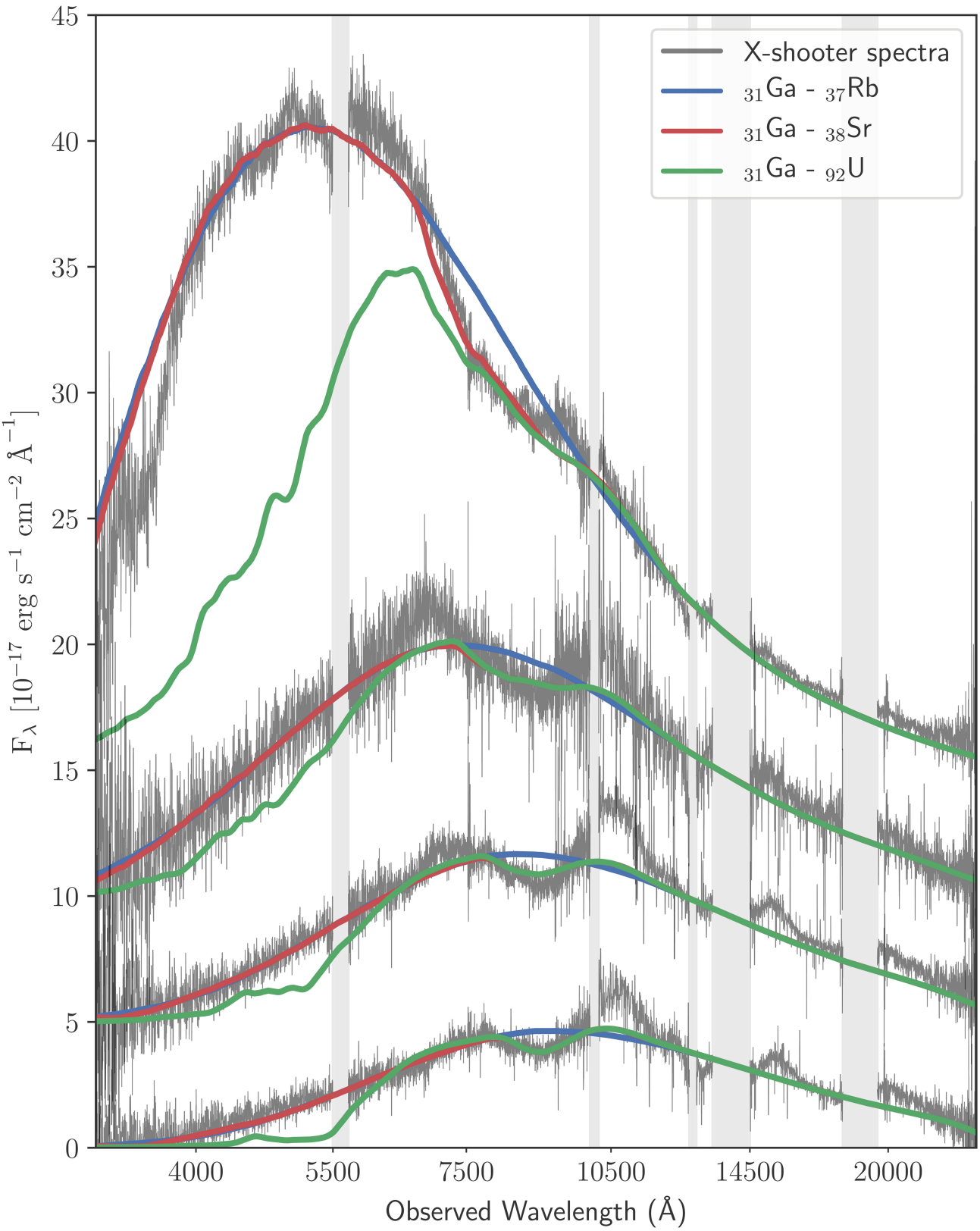

**Extended Data Fig. 3 | Thermal transmission spectra for r-process elements plotted individually.** The spectra are based on the lines formed in a gas in local thermal equilibrium. The abundances of elements are scaled to the solar r-process and the spectra are velocity broadened, blueshifted and normalized as in Fig. 3. The spectrum derived from the total solar r-process abundance mix is plotted as a thick black line. The contributions from Sr clearly dominate at around 8,000 Å, with no substantial contribution from any other element.




**Extended Data Fig. 4 | Evolution of the ejecta expansion velocity.** The velocities were determined independently from the P Cygni absorption line widths (blue points) and the blackbody radius (red points). Uncertainties shown are 1$\sigma$. The correspondence between the two independent estimates is striking.




**Extended Data Fig. 5 | Comparison of the expansion opacities at modest optical depths for Sr and Ce.** This calculation shows the potential of Sr to dominate the opacity at around 1 μm at low optical depths. The opacities are based on LTE calculations for a gas at a temperature of 5,000 K, a mean local density of $8.4 \times 10^{-17}$ g cm$^{-3}$ for Sr or Ce, an electron density of $7.6 \times 10^{8}$ cm$^{-3}$, and a 1% atmospheric radius at 1.5 days after the explosion. Line lists used for Sr and Ce are from the Kurucz and VALD databases respectively.




**Extended Data Fig. 6 | Radiative transfer models from the first four epochs using the TARDIS code.** The blue line is the synthetic TARDIS spectrum using relative solar r-process abundances and including elements from $_{31}$Ga to $_{37}$Rb—that is, without Sr. The red line also includes $_{38}$Sr. The green line is a model including all elements from $_{31}$Ga to $_{92}$U. These models show that the spectra are well reproduced with elements around the first r-process abundance peak, specifically Sr.

Article

# Topologically enabled ultrahigh-*Q* guided resonances robust to out-of-plane scattering

https://doi.org/10.1038/s41586-019-1664-7

Received: 21 January 2019

Accepted: 26 August 2019

Published online: 23 October 2019

Jicheng Jin[1], Xuefan Yin[1], Liangfu Ni[1], Marin Soljačić[2], Bo Zhen[3] & Chao Peng[1,4]*

Because of their ability to confine light, optical resonators[1–3] are of great importance to science and technology, but their performance is often limited by out-of-plane-scattering losses caused by inevitable fabrication imperfections[4,5]. Here we theoretically propose and experimentally demonstrate a class of guided resonances in photonic crystal slabs, in which out-of-plane-scattering losses are strongly suppressed by their topological nature. These resonances arise when multiple bound states in the continuum—each carrying a topological charge[6]—merge in momentum space and enhance the quality factors *Q* of all nearby resonances in the same band. Using such resonances in the telecommunication regime, we experimentally achieve quality factors as high as $4.9\times10^5$—12 times higher than those obtained with standard designs—and this enhancement remains robust for all of our samples. Our work paves the way for future explorations of topological photonics in systems with open boundary conditions and for their application to the improvement of optoelectronic devices in photonic integrated circuits.

Topological defects[7] are ubiquitous in the natural world. Examples range from quantum vortices in superfluids to singular optical beams[8], which are characterized by the non-trivial winding patterns of system parameters (velocity, phase or polarization) in real space. Recently, it was found that unexpected topological defects can also emerge in the momentum space of a crystal and give rise to interesting physical consequences; one such example is the optical bound states in the continuum (BICs). BICs reside inside the continuous spectrum of extended states but counter-intuitively remain perfectly localized in space and their lifetimes are theoretically infinitely long. Since their initial proposal[9], BICs have been observed in a variety of wave systems, including photonic[10–15], phononic[16] and water waves[17]. Furthermore, they have been used to enhance various applications, such as lasers[13,18] and antennas[19,20], by providing an out-coupling channel through their surface-emitting nature. In photonic crystal slabs, it has been identified that their fundamental nature is topological; they are essentially topological defects in polarization directions defined in momentum space[6]. In practice[10,21], the quality factors of such BICs are often much lower than their theoretical prediction of infinity, limited to only about $10^4$. Aside from other contributing factors, such as material absorption or the finite size of samples, the main limiting factor of the *Q* value of BICs is scattering losses caused by fabrication imperfections or disorders—a common problem for many high-*Q* on-chip resonators[1,2,4,5].

Here we theoretically propose and experimentally demonstrate on-chip photonic resonances that are much less susceptible to out-of-plane-scattering losses than expected, owing to their unique topological features. We start by showing that resonances with ultrahigh *Q* can be achieved by merging multiple BICs. First, we consider a photonic crystal slab (Fig. 1a), in which a square lattice (periodicity $a = 519.25$ nm) of circular air holes (radius $r = 175$ nm) is patterned in silicon (thickness $h = 600$ nm). With the use of numerical simulations (using the COMSOL Multiphysics software), we focus on the transverse electric (TE) A band (red line), featuring nine BICs where *Q* diverges to infinity (Fig. 1b). The topological nature of the BICs can be understood from the corresponding far-field polarization plots (Fig. 1b, bottom panels), where each BIC appears as a topological defect (vortex) in the polarization long axes[6,19,20,22,23] characterized by an integer topological charge of ±1. Among these nine vortices, one is pinned at the centre of the Brillouin zone owing to symmetry, whereas the locations of the remaining eight can be controlled by varying system parameters such as the periodicity. For example, when *a* increases from 519.25 nm to 531.42 nm, the eight off-centre BICs move towards the centre until all nine of them merge into a single BIC with a charge of +1 when $a = 531.42$ nm. This single BIC persists when *a* further increases to 580 nm.

The topological configuration of BICs controls radiative losses of all nearby resonances. Specifically, *Q* is shown to decay quadratically ($Q\propto1/k^2$) with respect to the distance *k* (in momentum space) from a single isolated BIC with charge ±1; however, this scaling changes to $Q\propto1/k^6$ in the configuration in which all nine BICs merge (referred to as 'merging-BIC design' hereafter). A comparison between these two scenarios is shown in Fig. 1c, where the *Q* values in a merging-BIC design (red) are always orders of magnitude higher than those in an isolated-BIC design (blue) along all directions in *k* space, owing to their fundamentally different scaling properties. This difference in scaling originates from the asymptotic behaviour of $Q\propto1/[k(k+k_{BIC})(k-k_{BIC})]^2$ in the regime in which off-centre BICs at $\pm k_{BIC}$ and centred BIC coexist, as shown in Fig. 1c (grey lines). In the merging-BIC design, $k_{BIC}\rightarrow0$ and we get $Q\propto1/k^6$. Further discussion is provided in Supplementary Information sections I and II.

Although simulation results of infinitely large perfect photonic crystals reproduce radiative quality factors, real samples (schematically

[1]State Key Laboratory of Advanced Optical Communication Systems and Networks, Department of Electronics, Peking University, Beijing, China. [2]Department of Physics, Massachusetts Institute of Technology, Cambridge, MA, USA. [3]Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA, USA. [4]Nano-optoelectronics Frontier Center of the Ministry of Education, Beijing, China. *e-mail: pengchao@pku.edu.cn

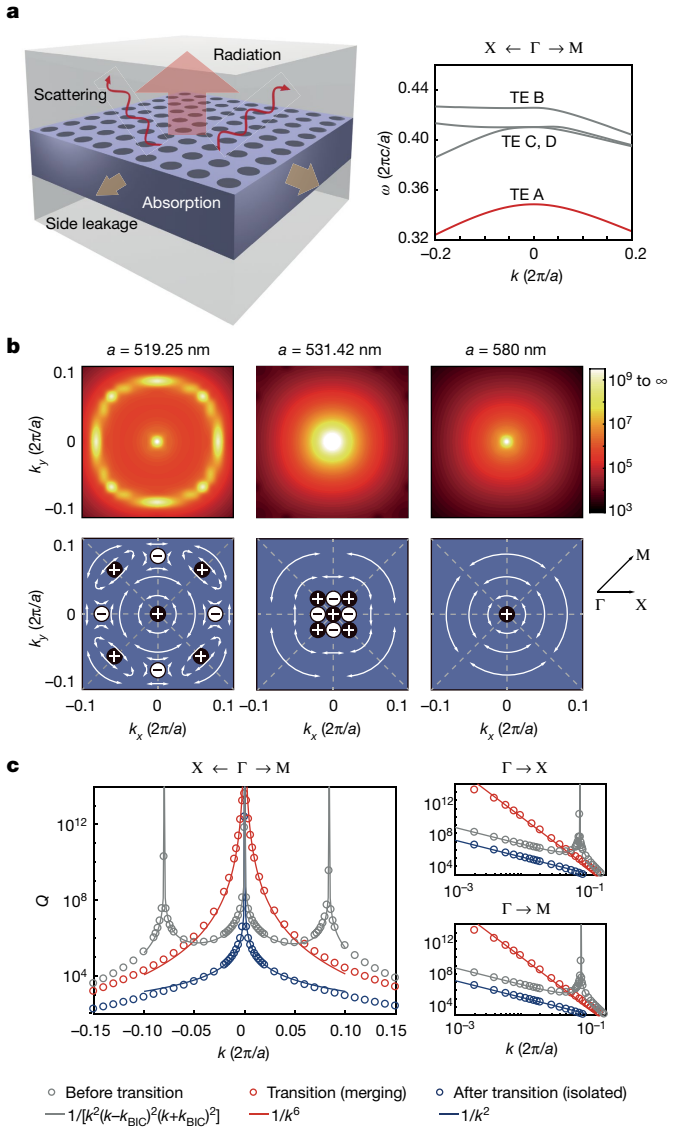



**Fig. 1 | Suppressing radiation losses by merging multiple topological charges. a**, Left, Schematic of a photonic crystal slab and the factors contributing to loss. Right, Simulated band structure. The TE A band is marked with a red line, and $\omega$ is the normalized resonance frequency. **b**, Multiple BICs appear on band TE A, where the radiative quality factor $Q$ diverges. Top, Simulated $Q$ for various values of the sample periodicity, $a$. Bottom, far-field polarization plots. When $a$ is tuned from 519.25 nm (left) to 580 nm (right), nine BICs at $k_{BIC}$ with topological charge ±1 merge into an isolated BIC with charge +1. **c**, Simulated $Q$ before ($a$ = 519.25 nm; grey) and after charges merge at the centre of the Brillouin zone ($a$ = 580 nm; blue). The transition ($a$ = 531.42 nm; red) corresponds to the merging-BIC configuration, which shows considerably higher quality factors than the isolated-BIC configuration (blue). This is caused by a change to a scaling rule of $Q \propto 1/k^6$, which is observed along both the Γ–X and Γ–M directions. All simulations used the finite-element method in COMSOL.

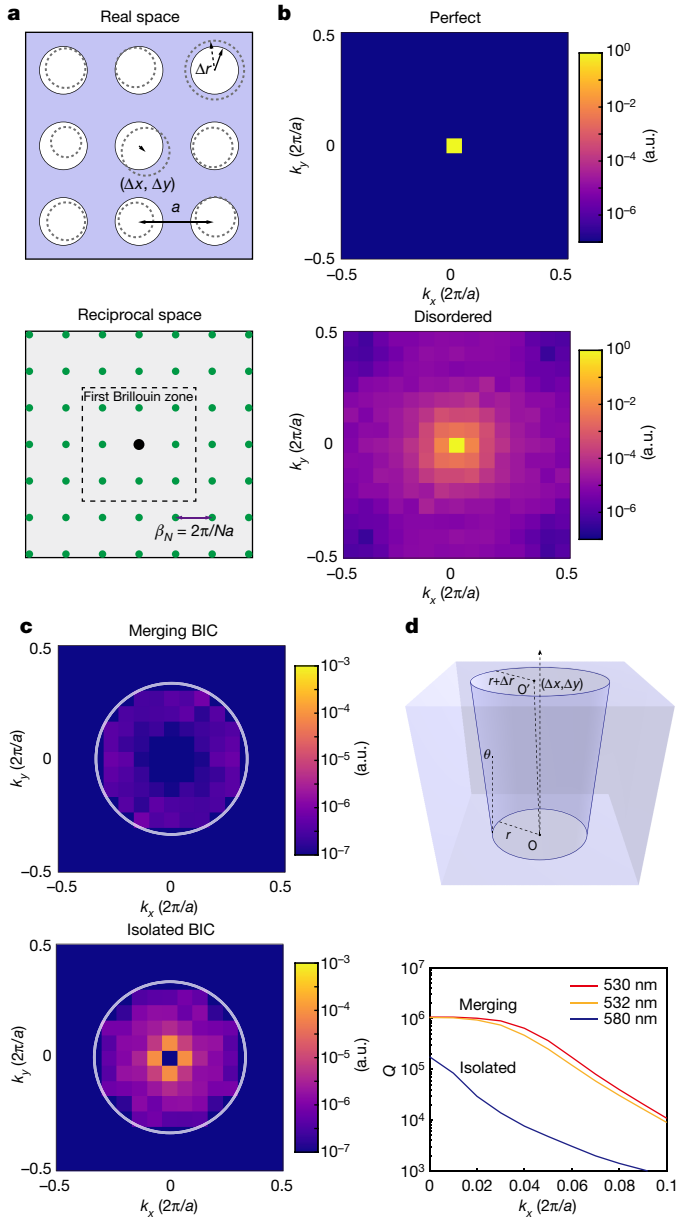



**Fig. 2 | Topological protection against scattering losses. a**, Schematic of a fabricated photonic crystal sample (solid lines; top), with disorder in the locations and radii of the holes (dashed lines; top). In reciprocal space, fractional orders of momentum (green dots; bottom) are introduced by the supercell. $N$ is the size of the supercell in increments of the periodicity $a$ and $\beta_N$ is the minimum step size of the fractional orders of momentum. **b**, Momentum–energy distribution of the highest-$Q$ mode in the merging-BIC design for a perfect (top) and a disordered (bottom) structure inside the first Brillouin zone. a.u., arbitrary units. **c**, Momentum–energy distribution of the radiation field in the disordered merging-BIC (top) and isolated-BIC (bottom) designs. The white circles represent the light cone—the region in the momentum space where guided resonance can couple to the radiation channel, as determined by the structure of the sample. The scattering loss is considerably lower in the merging-BIC sample than in the isolated-BIC sample. **d**, Schematic of an asymmetric hole (top) acting as a fabrication imperfection, and simulated $Q$ values near the centre of the Brillouin zone obtained by the application of disorder (bottom). All simulations were performed in a $15a \times 15a$ supercell ($N$ = 15).

shown in Fig. 2a) feature some major differences that determine the highest $Q$ achievable in practice. First, all samples are finite in size; their boundaries introduce fractional orders of the primitive reciprocal lattice in $k$ space (green dots in Fig. 2a; see Supplementary Information section III for details)[24]. Second, all fabricated samples exhibit disorder and imperfections with both long- and short-range correlations, allowing modes at different $k$ points to couple to each other. Because of these inevitable coupling terms, modes at different fractional momentum orders are hybridized and all of their loss channels become available to the final resonance[25].

The advantage of our merging-BIC design over an isolated-BIC design is confirmed by simulations (using the COMSOL Multiphysics software) of perturbed 15 × 15 photonic crystal supercells. In a perfect supercell structure without disorder, a BIC with infinite $Q$ remains at

the centre of the Brillouin zone (Fig. 2b, upper panel). For comparison, perturbations are applied to both the radii ($\Delta r$) and positions ($\Delta x$, $\Delta y$) of the holes according to the statistics that best describes our samples (Fig. 3). As expected, each mode in the disordered samples has multiple components in $k$ space. Furthermore, resonances in a disordered sample with a merging-BIC design have considerably lower radiation fields than those from an isolated-BIC design with the same disorder (Fig. 2c). This result agrees well with Fig. 1b, c: all modes contributing to the final resonance in the merging-BIC sample have much higher $Q$ values than those in the isolated-BIC case because resonances in the former are much more immune to out-of-plane scattering from disorder than in the latter. Finally, this enhancement of $Q$ is found to be robust across a range of $k$ values, as shown in Fig. 2d (bottom panel), by assuming asymmetric holes to represent typical fabrication errors before applying the disorder (see Supplementary Information sections IV and V for details).

To verify our theoretical findings, we fabricate photonic crystal samples with both merging-BIC and isolated-BIC designs using the same electron-beam lithography and inductively coupled plasma etching processes on a 600-nm-thick silicon-on-insulator wafer (see Methods for details). The underlying $SiO_2$ layer is then removed to restore the up–down mirror symmetry required for tunable BICs[6,10]. Alternatively, one may use refractive-index-matching liquid or deposition layers instead. The samples are about 250 × 250 μm$^2$ in size. The periodicity is varied from 530 nm to 580 nm to sample designs with merging and isolated BICs. From the scanning electron microscope images of the samples (Fig. 3a, b), the standard deviations of the hole locations and radii are estimated to be about 5 nm, which is used in the simulations discussed earlier.

A schematic of the experimental setup is shown in Fig. 3c. A tunable telecommunication laser with light in the C+L band is first sent through an X-polarizer before the light is focused by a lens (L1) onto the back focal plane of an infinity-corrected objective lens. The incident angle of the laser on the sample is thus controlled by moving L1 on the $x$–$y$ plane. Using this confocal setup, reflected and scattered light are also collected by the same objective; they are then magnified 1.67 times through a relay 4f system and imaged on a camera. A Y-polarizer is used to block reflected light (X-polarized) while allowing scattered light to pass (see Methods and Supplementary Information section VI). Under the on-resonance coupling condition, where the photonic crystal sample supports a resonance at the same wavelength as the incident light at that incident angle, isofrequency contours are observed on the camera, similarly to previously reported results[22,26]. Three examples of isofrequency contours are schematically shown in Fig. 4a as dashed lines.

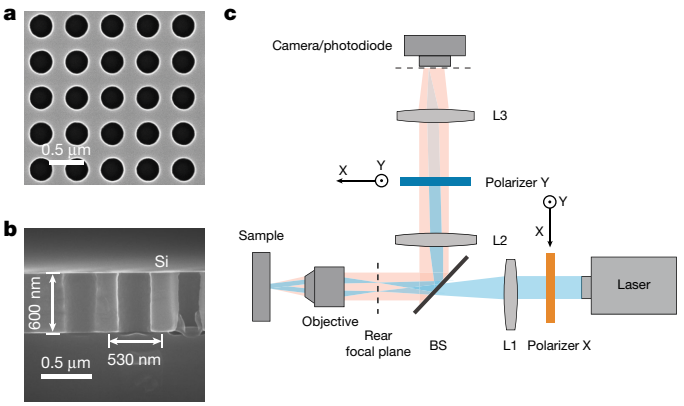



**Fig. 3 | Experimental setup. a**, **b**, Scanning electron microscope images of the fabricated photonic crystal sample; top (**a**) and side (**b**) view. **c**, Schematic of the measurement setup. The blue lines represent the incident light and its direct reflection. The light red region denotes radiation losses induced by scattering from disorder. L, lens. BS, beamsplitter.

The $Q$ values of resonances at different $k$ points are further characterized using scattered light. Specifically, a movable pinhole (not shown in Fig. 3c) is placed on the image plane of the objective's rear focal plane to specify a $k$ point. A photodiode connected to a lock-in amplifier is placed behind the pinhole to record the intensity of the scattered light as a function of the wavelength of the tunable laser. As shown in Fig. 4a, when different $k$ points are selected by the pinhole (X, Y and Z), different scattering spectra are obtained, all exhibiting symmetric Lorentzian features. Similar scattering phenomena have been observed before[26] and can be understood as follows: the intensity of scattered light is governed by the spectral density of states of the sample at that $k$ point, and is described by a Lorentzian function centred at the resonance frequency, with linewidth determined by the $Q$ value of the resonance (see Supplementary Information section VII for details).

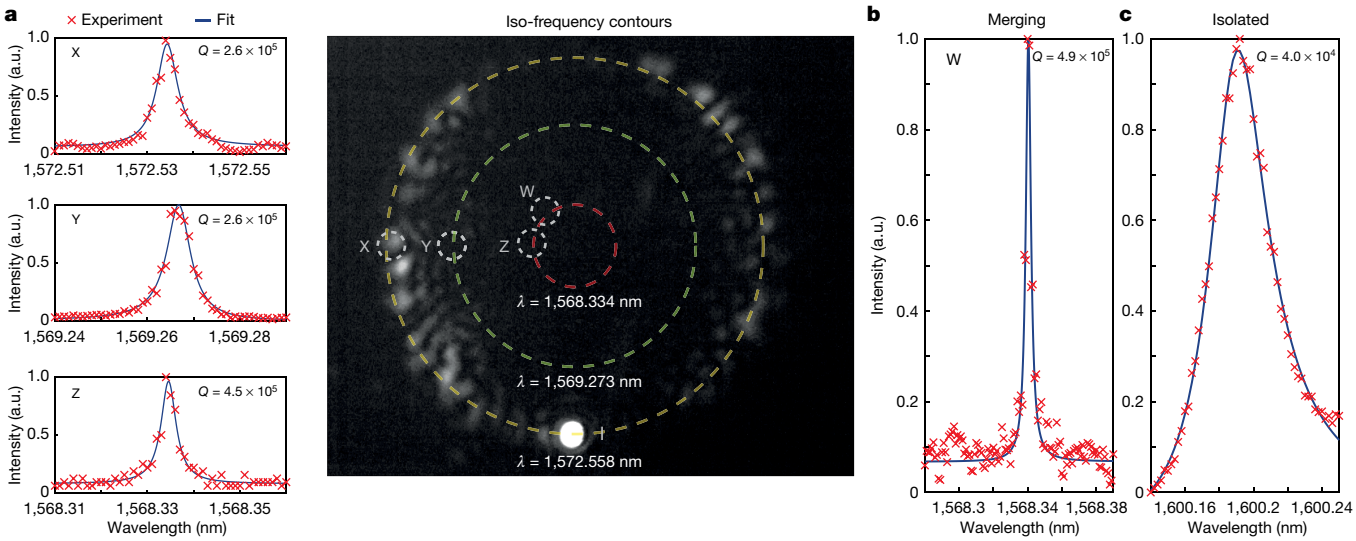



**Fig. 4 | Experimental results. a**, Isofrequency contours of the sample at different wavelengths are observed on the camera (right). Three examples are shown as dashed lines. The scattered-light intensity at different points in momentum space (X, Y, Z; left) is further characterized using a photodiode, and is fitted by symmetric Lorentzians as a function of incident wavelength. The linewidth is determined by the $Q$ value of the underlying resonance. **b**, **c**, The highest $Q$ observed in the merging-BIC sample is $4.9 \times 10^5$ at point W (**b**), which is more than an order of magnitude larger than that of an isolated-BIC sample constructed with the same fabrication process ($Q = 4.0 \times 10^4$; **c**).

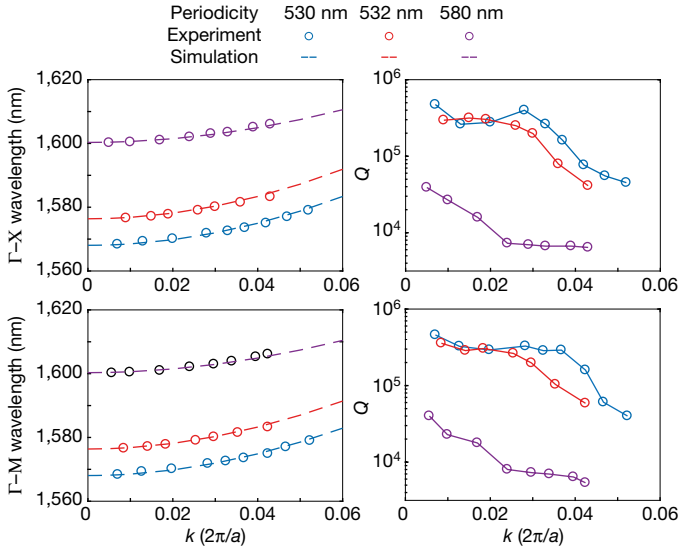



**Fig. 5 | 12-fold enhancement of *Q* via topological protection. a**, Dispersion of resonances measured at different points in momentum space (circles), showing good agreement with FEM simulation predictions (dashed lines) along both the Γ–X (top) and Γ–M (bottom) directions. **b**, An enhancement of *Q* by a factor of more than 10 is observed over a wide range in momentum space for the merging-BIC samples (red and blue) compared to the isolated-BIC sample (purple) owing to topological protection.

The *Q* values of the resonances are extracted by numerically fitting the scattering spectra with Lorentzian functions. As shown in Fig. 4a, *Q* increases from $2.6\times10^5$ to $4.5\times10^5$ as the observing point moves closer to the centre of the Brillouin zone from X to Z. This agrees well with the simulation results in Fig. 1. The highest *Q* observed in the merging-BIC sample is $4.9\times10^5$ at point W (Fig. 4b). In comparison, the highest *Q* observed in the isolated-BIC sample—fabricated on the same wafer through the same processes as the merging-BIC sample, but with different structural parameters—is limited to only $4\times10^4$, more than an order of magnitude lower (Fig. 4c). This confirms our simulation results in Fig. 2, which indicate that engineering the topological configurations of BICs can substantially suppress scattering losses. Furthermore, this over-ten-fold enhancement of the quality factor is observed to be robust: not only does it appear over a wide range in *k* space, as shown in Fig. 5, but a similar level of enhancement is also observed in all merging-BIC samples that we fabricated (see Supplementary Information section VIII for details).

Topological photonics[27–29] has found tremendous success in suppressing in-plane back-scattering losses by breaking reciprocity. Here we use topology to solve a different class of problems, by suppressing out-of-plane-scattering losses in a reciprocal system. By merging multiple topological charges carried by BICs, we experimentally demonstrate photonic crystal resonances with record-high quality factors of $Q=4.9\times10^5$, more than an order of magnitude higher than those of ordinary designs. These ultrahigh-*Q* resonances are potentially useful for chemical or biological sensing[30] and large-area laser applications[18]. Furthermore, our high-*Q* resonances are observed to be robust against fabrication imperfections, and can help to improve the performance of optoelectronic devices using concepts from topological photonics.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1664-7.


1. Biberman, A., Shaw, M. J., Timurdogan, E., Wright, J. B. & Watts, M. R. Ultralow-loss silicon ring resonators. *Opt. Lett.* **37**, 4236–4238 (2012).
2. Hossein-Zadeh, M. & Vahala, K. J. Free ultra-high-q microtoroid: a tool for designing photonic devices. *Opt. Express* **15**, 166–175 (2007).
3. Akahane, Y., Asano, T., Song, B.-S. & Noda, S. High-*Q* photonic nanocavity in a two-dimensional photonic crystal. *Nature* **425**, 944–947 (2003).
4. Minkov, M., Dharanipathy, U. P., Houdré, R. & Savona, V. Statistics of the disorder-induced losses of high-*Q* photonic crystal cavities. *Opt. Express* **21**, 28233–28245 (2013).
5. Ishizaki, K., Okano, M. & Noda, S. Numerical investigation of emission in finite-sized, three-dimensional photonic crystals with structural fluctuations. *J. Opt. Soc. Am. B* **26**, 1157–1161 (2009).
6. Zhen, B., Hsu, C. W., Lu, L., Stone, A. D. & Soljačić, M. Topological nature of optical bound states in the continuum. *Phys. Rev. Lett.* **113**, 257401 (2014).
7. Mermin, N. D. The topological theory of defects in ordered media. *Rev. Mod. Phys.* **51**, 591–648 (1979).
8. Gbur, G. J. *Singular Optics* (CRC Press, 2016).
9. von Neumann, J. & Wigner, E. Über merkwürdige diskrete Eigenwerte. *Phys. Z.* **30**, 465–467 (1929).
10. Hsu, C. W. et al. Observation of trapped light within the radiation continuum. *Nature* **499**, 188–191 (2013).
11. Plotnik, Y. et al. Experimental observation of optical bound states in the continuum. *Phys. Rev. Lett.* **107**, 183901 (2011).
12. Fan, S. & Joannopoulos, J. D. Analysis of guided resonances in photonic crystal slabs. *Phys. Rev. B* **65**, 235112 (2002).
13. Kodigala, A. et al. Lasing action from photonic bound states in continuum. *Nature* **541**, 196–199 (2017).
14. Gomis-Bresco, J., Artigas, D. & Torner, L. Anisotropy-induced photonic bound states in the continuum. *Nat. Photon.* **11**, 232–236 (2017).
15. Molina, M. I., Miroshnichenko, A. E. & Kivshar, Y. S. Surface bound states in the continuum. *Phys. Rev. Lett.* **108**, 070401 (2012).
16. Lim, T. C. & Farnell, G. W. Character of pseudo surface waves on anisotropic crystals. *J. Acoust. Soc. Am.* **45**, 845–851 (1969).
17. Cobelli, P. J., Pagneux, V., Maurel, A. & Petitjeans, P. Experimental observation of trapped modes in a water wave channel. *EPL* **88**, 20006 (2009).
18. Hirose, K. et al. Watt-class high-power, high-beam-quality photonic-crystal lasers. *Nat. Photon.* **8**, 406–411 (2014).
19. Kitamura, K., Sakai, K., Takayama, N., Nishimoto, M. & Noda, S. Focusing properties of vector vortex beams emitted by photonic-crystal lasers. *Opt. Lett.* **37**, 2421–2423 (2012).
20. Iwahashi, S. et al. Higher-order vector beams produced by photonic-crystal lasers. *Opt. Express* **19**, 11963–11968 (2011).
21. Lee, J., Zhen, B., Chua, S.-L., Shapira, O. & Soljačić, M. Fabricating centimeter-scale high quality factor two-dimensional periodic photonic crystal slabs. *Opt. Express* **22**, 3724–3731 (2014).
22. Zhou, H. et al. Observation of bulk Fermi arc and polarization half charge from paired exceptional points. *Science* **359**, eaap9859 (2018).
23. Zhang, Y. et al. Observation of polarization vortices in momentum space. *Phys. Rev. Lett.* **120**, 186103 (2018).
24. Liang, Y., Peng, C., Sakai, K., Iwahashi, S. & Noda, S. Three-dimensional coupled-wave analysis for square-lattice photonic crystal surface emitting lasers with transverse-electric polarization: finite-size effects. *Opt. Express* **20**, 15945–15961 (2012).
25. Ni, L., Jin, J., Peng, C. & Li, Z. Analytical and statistical investigation on structural fluctuations induced radiation in photonic crystal slabs. *Opt. Express* **25**, 5580–5593 (2017).
26. Regan, E. C. et al. Direct imaging of isofrequency contours in photonic structures. *Sci. Adv.* **2**, e1601591 (2016).
27. Lu, L., Joannopoulos, J. D. & Soljačić, M. Topological photonics. *Nat. Photon.* **8**, 821–829 (2014).
28. Ozawa, T. et al. Topological photonics. *Rev. Mod. Phys.* **91**, 015006 (2019).
29. Wang, Z., Chong, Y., Joannopoulos, J. D. & Soljačić, M. Observation of unidirectional backscattering-immune topological electromagnetic states. *Nature* **461**, 772–775 (2009).
30. Luchansky, M. S. & Bailey, R. C. High-*Q* optical sensors for chemical and biological analysis. *Anal. Chem.* **84**, 793–821 (2012).

# Methods

## Sample fabrication

The sample was fabricated on a silicon-on-insulator wafer with electron-beam lithography (EBL), followed by inductively coupled plasma etching. For EBL, the silicon-on-insulator wafer was spin-coated with a 330-nm-thick layer of ZEP520A photo-resist before being exposed to EBL (JBX-6300FS) at a beam current of 400 pA and a field size of 500 μm. The sample was then etched with ICP (Oxford Plasmapro Estrelas 100) using a mixture of $SF_6$ and $C_4F_8$. After etching, the resist was removed with *N*-methyl-2-pyrrolidone and the buried oxide layer was removed using 49% HF.

## Measurement system

The incident light source was a tunable C+L-band telecommunication laser (Santec TSL-550), which was sent through a chopper for lock-in detection. A pinhole with diameter of 500 μm was placed on the Fourier plane to select the desired wavevectors. Light scattered through the pinhole was collected by a photodiode (PDA10DT-EC), which was connected to a lock-in amplifier (SRS SR830). A flip mirror was used to switch between the camera that was used to image isofrequency contours and the photodiode. Besides characterizing far-field radiation patterns, the setup could also take near-field images of the sample if another lens was inserted into the optical path.

## Data availability

The data that support the plots in this paper and other findings of this study are available from the corresponding author upon request.

**Acknowledgements** We thank C. W. Hsu for discussions. We also thank L. Zhang from Tsinghua University and W. Liu from the Institute of Semiconductors for assistance in the preparation of the device. C.P. was supported by the National Natural Science Foundation of China under grant number 61575002. B.Z. was supported by the US Air Force Office of Scientific Research under award number FA9550-18-1-0133, by the US National Science Foundation through grant DMR-1838412 and by the Army Research Office under grant number W911NF-19-1-0087. The research was sponsored in part by the Army Research Office under Cooperative Agreement number W911NF-18-2-0048.

**Author contributions** J.J., X.Y., L.N. and C.P. conceived the idea. J.J., X.Y., C.P., B.Z. and M.S. performed the theoretical study. J.J., X.Y., L.N. and C.P. performed the analytical calculations and numerical simulations. J.J. and X.Y. conducted the experiments and analysed the data. J.J., C.P. and B.Z. wrote the manuscript, with input from all authors. C.P. supervised the research. All authors discussed the results.

**Competing interests** The authors declare no competing interests.

**Additional information**
**Supplementary information** is available for this paper at https://doi.org/10.1038/s41586-019-1664-7.
**Correspondence and requests for materials** should be addressed to C.P.
**Reprints and permissions information** is available at http://www.nature.com/reprints.

Article

# Quantum supremacy using a programmable superconducting processor

https://doi.org/10.1038/s41586-019-1666-5

Received: 22 July 2019

Accepted: 20 September 2019

Published online: 23 October 2019

Frank Arute[1], Kunal Arya[1], Ryan Babbush[1], Dave Bacon[1], Joseph C. Bardin[1,2], Rami Barends[1], Rupak Biswas[3], Sergio Boixo[1], Fernando G. S. L. Brandao[1,4], David A. Buell[1], Brian Burkett[1], Yu Chen[1], Zijun Chen[1], Ben Chiaro[5], Roberto Collins[1], William Courtney[1], Andrew Dunsworth[1], Edward Farhi[1], Brooks Foxen[1,5], Austin Fowler[1], Craig Gidney[1], Marissa Giustina[1], Rob Graff[1], Keith Guerin[1], Steve Habegger[1], Matthew P. Harrigan[1], Michael J. Hartmann[1,6], Alan Ho[1], Markus Hoffmann[1], Trent Huang[1], Travis S. Humble[7], Sergei V. Isakov[1], Evan Jeffrey[1], Zhang Jiang[1], Dvir Kafri[1], Kostyantyn Kechedzhi[1], Julian Kelly[1], Paul V. Klimov[1], Sergey Knysh[1], Alexander Korotkov[1,8], Fedor Kostritsa[1], David Landhuis[1], Mike Lindmark[1], Erik Lucero[1], Dmitry Lyakh[9], Salvatore Mandrà[3,10], Jarrod R. McClean[1], Matthew McEwen[5], Anthony Megrant[1], Xiao Mi[1], Kristel Michielsen[11,12], Masoud Mohseni[1], Josh Mutus[1], Ofer Naaman[1], Matthew Neeley[1], Charles Neill[1], Murphy Yuezhen Niu[1], Eric Ostby[1], Andre Petukhov[1], John C. Platt[1], Chris Quintana[1], Eleanor G. Rieffel[3], Pedram Roushan[1], Nicholas C. Rubin[1], Daniel Sank[1], Kevin J. Satzinger[1], Vadim Smelyanskiy[1], Kevin J. Sung[1,13], Matthew D. Trevithick[1], Amit Vainsencher[1], Benjamin Villalonga[1,14], Theodore White[1], Z. Jamie Yao[1], Ping Yeh[1], Adam Zalcman[1], Hartmut Neven[1] & John M. Martinis[1,5]*

The promise of quantum computers is that certain computational tasks might be executed exponentially faster on a quantum processor than on a classical processor[1]. A fundamental challenge is to build a high-fidelity processor capable of running quantum algorithms in an exponentially large computational space. Here we report the use of a processor with programmable superconducting qubits[2–7] to create quantum states on 53 qubits, corresponding to a computational state-space of dimension $2^{53}$ (about $10^{16}$). Measurements from repeated experiments sample the resulting probability distribution, which we verify using classical simulations. Our Sycamore processor takes about 200 seconds to sample one instance of a quantum circuit a million times—our benchmarks currently indicate that the equivalent task for a state-of-the-art classical supercomputer would take approximately 10,000 years. This dramatic increase in speed compared to all known classical algorithms is an experimental realization of quantum supremacy[8–14] for this specific computational task, heralding a much-anticipated computing paradigm.

In the early 1980s, Richard Feynman proposed that a quantum computer would be an effective tool with which to solve problems in physics and chemistry, given that it is exponentially costly to simulate large quantum systems with classical computers[1]. Realizing Feynman's vision poses substantial experimental and theoretical challenges. First, can a quantum system be engineered to perform a computation in a large enough computational (Hilbert) space and with a low enough error rate to provide a quantum speedup? Second, can we formulate a problem that is hard for a classical computer but easy for a quantum computer? By computing such a benchmark task on our superconducting qubit processor, we tackle both questions. Our experiment achieves quantum supremacy, a milestone on the path to full-scale quantum computing[8–14].

In reaching this milestone, we show that quantum speedup is achievable in a real-world system and is not precluded by any hidden physical laws. Quantum supremacy also heralds the era of noisy intermediate-scale quantum (NISQ) technologies[15]. The benchmark task we demonstrate has an immediate application in generating certifiable random numbers (S. Aaronson, manuscript in preparation); other initial uses for this new computational capability may include optimization[16,17], machine learning[18–21], materials science and chemistry[22–24]. However, realizing the full promise of quantum computing (using Shor's algorithm for factoring, for example) still requires technical leaps to engineer fault-tolerant logical qubits[25–29].

To achieve quantum supremacy, we made a number of technical advances which also pave the way towards error correction. We

[1]Google AI Quantum, Mountain View, CA, USA. [2]Department of Electrical and Computer Engineering, University of Massachusetts Amherst, Amherst, MA, USA. [3]Quantum Artificial Intelligence Laboratory (QuAIL), NASA Ames Research Center, Moffett Field, CA, USA. [4]Institute for Quantum Information and Matter, Caltech, Pasadena, CA, USA. [5]Department of Physics, University of California, Santa Barbara, CA, USA. [6]Friedrich-Alexander University Erlangen-Nürnberg (FAU), Department of Physics, Erlangen, Germany. [7]Quantum Computing Institute, Oak Ridge National Laboratory, Oak Ridge, TN, USA. [8]Department of Electrical and Computer Engineering, University of California, Riverside, CA, USA. [9]Scientific Computing, Oak Ridge Leadership Computing, Oak Ridge National Laboratory, Oak Ridge, TN, USA. [10]Stinger Ghaffarian Technologies Inc., Greenbelt, MD, USA. [11]Institute for Advanced Simulation, Jülich Supercomputing Centre, Forschungszentrum Jülich, Jülich, Germany. [12]RWTH Aachen University, Aachen, Germany. [13]Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI, USA. [14]Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL, USA. *e-mail: jmartinis@google.com

developed fast, high-fidelity gates that can be executed simultaneously across a two-dimensional qubit array. We calibrated and benchmarked the processor at both the component and system level using a powerful new tool: cross-entropy benchmarking[11]. Finally, we used component-level fidelities to accurately predict the performance of the whole system, further showing that quantum information behaves as expected when scaling to large systems.

## A suitable computational task

To demonstrate quantum supremacy, we compare our quantum processor against state-of-the-art classical computers in the task of sampling the output of a pseudo-random quantum circuit[11,13,14]. Random circuits are a suitable choice for benchmarking because they do not possess structure and therefore allow for limited guarantees of computational hardness[10–12]. We design the circuits to entangle a set of quantum bits (qubits) by repeated application of single-qubit and two-qubit logical operations. Sampling the quantum circuit's output produces a set of bitstrings, for example {0000101, 1011100, ...}. Owing to quantum interference, the probability distribution of the bitstrings resembles a speckled intensity pattern produced by light interference in laser scatter, such that some bitstrings are much more likely to occur than others. Classically computing this probability distribution becomes exponentially more difficult as the number of qubits (width) and number of gate cycles (depth) grow.

We verify that the quantum processor is working properly using a method called cross-entropy benchmarking[11,12,14], which compares how often each bitstring is observed experimentally with its corresponding ideal probability computed via simulation on a classical computer. For a given circuit, we collect the measured bitstrings $\{x_i\}$ and compute the linear cross-entropy benchmarking fidelity[11,13,14] (see also Supplementary Information), which is the mean of the simulated probabilities of the bitstrings we measured:

$$\mathcal{F}_{\mathrm{XEB}} = 2^n \langle P(x_i) \rangle_i - 1 \qquad (1)$$

where $n$ is the number of qubits, $P(x_i)$ is the probability of bitstring $x_i$ computed for the ideal quantum circuit, and the average is over the observed bitstrings. Intuitively, $\mathcal{F}_{\mathrm{XEB}}$ is correlated with how often we sample high-probability bitstrings. When there are no errors in the quantum circuit, the distribution of probabilities is exponential (see Supplementary Information), and sampling from this distribution will produce $\mathcal{F}_{\mathrm{XEB}} = 1$. On the other hand, sampling from the uniform distribution will give $\langle P(x_i) \rangle_i = 1/2^n$ and produce $\mathcal{F}_{\mathrm{XEB}} = 0$. Values of $\mathcal{F}_{\mathrm{XEB}}$ between 0 and 1 correspond to the probability that no error has occurred while running the circuit. The probabilities $P(x_i)$ must be obtained from classically simulating the quantum circuit, and thus computing $\mathcal{F}_{\mathrm{XEB}}$ is intractable in the regime of quantum supremacy. However, with certain circuit simplifications, we can obtain quantitative fidelity estimates of a fully operating processor running wide and deep quantum circuits.

Our goal is to achieve a high enough $\mathcal{F}_{\mathrm{XEB}}$ for a circuit with sufficient width and depth such that the classical computing cost is prohibitively large. This is a difficult task because our logic gates are imperfect and the quantum states we intend to create are sensitive to errors. A single bit or phase flip over the course of the algorithm will completely shuffle the speckle pattern and result in close to zero fidelity[11] (see also Supplementary Information). Therefore, in order to claim quantum supremacy we need a quantum processor that executes the program with sufficiently low error rates.

## Building a high-fidelity processor

We designed a quantum processor named 'Sycamore' which consists of a two-dimensional array of 54 transmon qubits, where each qubit is tunably coupled to four nearest neighbours, in a rectangular lattice. The

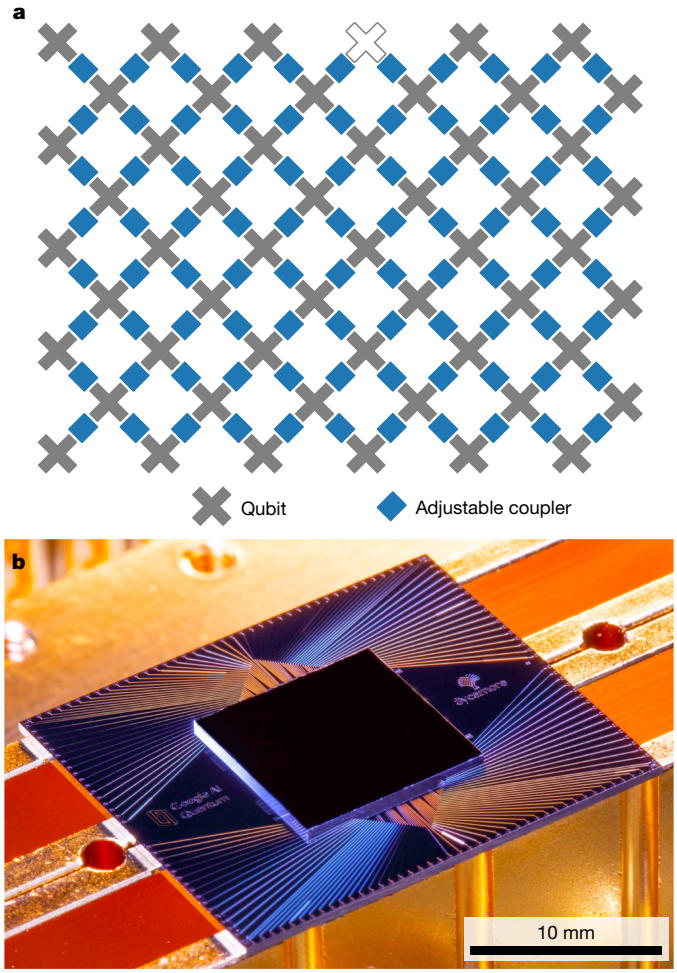



**Fig. 1 | The Sycamore processor. a**, Layout of processor, showing a rectangular array of 54 qubits (grey), each connected to its four nearest neighbours with couplers (blue). The inoperable qubit is outlined. **b**, Photograph of the Sycamore chip.

connectivity was chosen to be forward-compatible with error correction using the surface code[26]. A key systems engineering advance of this device is achieving high-fidelity single- and two-qubit operations, not just in isolation but also while performing a realistic computation with simultaneous gate operations on many qubits. We discuss the highlights below; see also the Supplementary Information.

In a superconducting circuit, conduction electrons condense into a macroscopic quantum state, such that currents and voltages behave quantum mechanically[2,30]. Our processor uses transmon qubits[6], which can be thought of as nonlinear superconducting resonators at 5–7 GHz. The qubit is encoded as the two lowest quantum eigenstates of the resonant circuit. Each transmon has two controls: a microwave drive to excite the qubit, and a magnetic flux control to tune the frequency. Each qubit is connected to a linear resonator used to read out the qubit state[5]. As shown in Fig. 1, each qubit is also connected to its neighbouring qubits using a new adjustable coupler[31,32]. Our coupler design allows us to quickly tune the qubit–qubit coupling from completely off to 40 MHz. One qubit did not function properly, so the device uses 53 qubits and 86 couplers.

The processor is fabricated using aluminium for metallization and Josephson junctions, and indium for bump-bonds between two silicon wafers. The chip is wire-bonded to a superconducting circuit board and cooled to below 20 mK in a dilution refrigerator to reduce ambient thermal energy to well below the qubit energy. The processor is connected through filters and attenuators to room-temperature electronics,

which synthesize the control signals. The state of all qubits can be read simultaneously by using a frequency-multiplexing technique[33,34]. We use two stages of cryogenic amplifiers to boost the signal, which is digitized (8 bits at 1 GHz) and demultiplexed digitally at room temperature. In total, we orchestrate 277 digital-to-analog converters (14 bits at 1 GHz) for complete control of the quantum processor.

We execute single-qubit gates by driving 25-ns microwave pulses resonant with the qubit frequency while the qubit–qubit coupling is turned off. The pulses are shaped to minimize transitions to higher transmon states[35]. Gate performance varies strongly with frequency owing to two-level-system defects[36,37], stray microwave modes, coupling to control lines and the readout resonator, residual stray coupling between qubits, flux noise and pulse distortions. We therefore optimize the single-qubit operation frequencies to mitigate these error mechanisms.

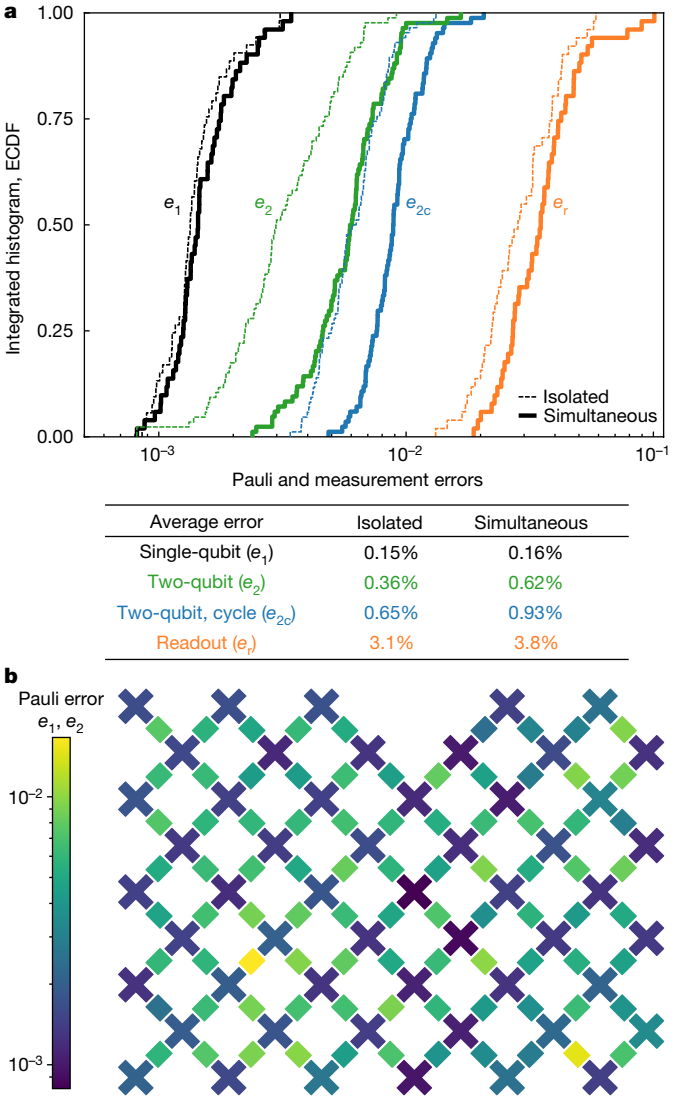
We benchmark single-qubit gate performance by using the cross-entropy benchmarking protocol described above, reduced to the single-qubit level ($n = 1$), to measure the probability of an error occurring during a single-qubit gate. On each qubit, we apply a variable number $m$ of randomly selected gates and measure $\mathcal{F}_{\mathrm{XEB}}$ averaged over many sequences; as $m$ increases, errors accumulate and average $\mathcal{F}_{\mathrm{XEB}}$ decays. We model this decay by $[1-e_1/(1-1/D^2)]^m$ where $e_1$ is the Pauli error probability. The state (Hilbert) space dimension term, $D = 2^n$, which equals 2 for this case, corrects for the depolarizing model where states with errors partially overlap with the ideal state. This procedure is similar to the more typical technique of randomized benchmarking[27,38,39], but supports non-Clifford-gate sets[40] and can separate out decoherence error from coherent control error. We then repeat the experiment with all qubits executing single-qubit gates simultaneously (Fig. 2), which shows only a small increase in the error probabilities, demonstrating that our device has low microwave crosstalk.

We perform two-qubit iSWAP-like entangling gates by bringing neighbouring qubits on-resonance and turning on a 20-MHz coupling for 12 ns, which allows the qubits to swap excitations. During this time, the qubits also experience a controlled-phase (CZ) interaction, which originates from the higher levels of the transmon. The two-qubit gate frequency trajectories of each pair of qubits are optimized to mitigate the same error mechanisms considered in optimizing single-qubit operation frequencies.

To characterize and benchmark the two-qubit gates, we run two-qubit circuits with $m$ cycles, where each cycle contains a randomly chosen single-qubit gate on each of the two qubits followed by a fixed two-qubit gate. We learn the parameters of the two-qubit unitary (such as the amount of iSWAP and CZ interaction) by using $\mathcal{F}_{\mathrm{XEB}}$ as a cost function. After this optimization, we extract the per-cycle error $e_{2c}$ from the decay of $\mathcal{F}_{\mathrm{XEB}}$ with $m$, and isolate the two-qubit error $e_2$ by subtracting the two single-qubit errors $e_1$. We find an average $e_2$ of 0.36%. Additionally, we repeat the same procedure while simultaneously running two-qubit circuits for the entire array. After updating the unitary parameters to account for effects such as dispersive shifts and crosstalk, we find an average $e_2$ of 0.62%.

For the full experiment, we generate quantum circuits using the two-qubit unitaries measured for each pair during simultaneous operation, rather than a standard gate for all pairs. The typical two-qubit gate is a full iSWAP with 1/6th of a full CZ. Using individually calibrated gates in no way limits the universality of the demonstration. One can compose, for example, controlled-NOT (CNOT) gates from 1-qubit gates and two of the unique 2-qubit gates of any given pair. The implementation of high-fidelity 'textbook gates' natively, such as CZ or $\sqrt{\mathrm{iSWAP}}$, is work in progress.

Finally, we benchmark qubit readout using standard dispersive measurement[41]. Measurement errors averaged over the 0 and 1 states are shown in Fig. 2a. We have also measured the error when operating all qubits simultaneously, by randomly preparing each qubit in the 0 or 1 state and then measuring all qubits for the probability of the correct result. We find that simultaneous readout incurs only a modest increase in per-qubit measurement errors.




| Average error | Isolated | Simultaneous |
|---|---|---|
| Single-qubit ($e_1$) | 0.15% | 0.16% |
| Two-qubit ($e_2$) | 0.36% | 0.62% |
| Two-qubit, cycle ($e_{2c}$) | 0.65% | 0.93% |
| Readout ($e_r$) | 3.1% | 3.8% |

**Fig. 2 | System-wide Pauli and measurement errors. a**, Integrated histogram (empirical cumulative distribution function, ECDF) of Pauli errors (black, green, blue) and readout errors (orange), measured on qubits in isolation (dotted lines) and when operating all qubits simultaneously (solid). The median of each distribution occurs at 0.50 on the vertical axis. Average (mean) values are shown below. **b**, Heat map showing single- and two-qubit Pauli errors $e_1$ (crosses) and $e_2$ (bars) positioned in the layout of the processor. Values are shown for all qubits operating simultaneously.

Having found the error rates of the individual gates and readout, we can model the fidelity of a quantum circuit as the product of the probabilities of error-free operation of all gates and measurements. Our largest random quantum circuits have 53 qubits, 1,113 single-qubit gates, 430 two-qubit gates, and a measurement on each qubit, for which we predict a total fidelity of 0.2%. This fidelity should be resolvable with a few million measurements, since the uncertainty on $\mathcal{F}_{\mathrm{XEB}}$ is $1/\sqrt{N_s}$, where $N_s$ is the number of samples. Our model assumes that entangling larger and larger systems does not introduce additional error sources beyond the errors we measure at the single- and two-qubit level. In the next section we will see how well this hypothesis holds up.

## Fidelity estimation in the supremacy regime

The gate sequence for our pseudo-random quantum circuit generation is shown in Fig. 3. One cycle of the algorithm consists of applying

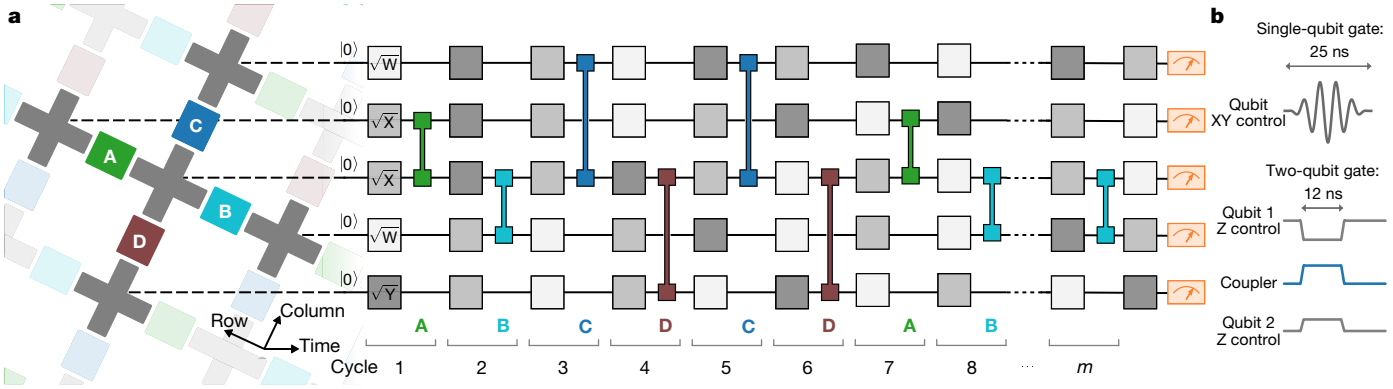



**Fig. 3 | Control operations for the quantum supremacy circuits. a**, Example quantum circuit instance used in our experiment. Every cycle includes a layer each of single- and two-qubit gates. The single-qubit gates are chosen randomly from $\{\sqrt{X}, \sqrt{Y}, \sqrt{W}\}$, where $W = (X+Y)/\sqrt{2}$ and gates do not repeat sequentially. The sequence of two-qubit gates is chosen according to a tiling pattern, coupling each qubit sequentially to its four nearest-neighbour qubits. The couplers are divided into four subsets (ABCD), each of which is executed simultaneously across the entire array corresponding to shaded colours. Here we show an intractable sequence (repeat ABCDCDAB); we also use different coupler subsets along with a simplifiable sequence (repeat EFGHEFGH, not shown) that can be simulated on a classical computer. **b**, Waveform of control signals for single- and two-qubit gates.

single-qubit gates chosen randomly from $\{\sqrt{X}, \sqrt{Y}, \sqrt{W}\}$ on all qubits, followed by two-qubit gates on pairs of qubits. The sequences of gates which form the 'supremacy circuits' are designed to minimize the circuit depth required to create a highly entangled state, which is needed for computational complexity and classical hardness.

Although we cannot compute $\mathcal{F}_{\mathrm{XEB}}$ in the supremacy regime, we can estimate it using three variations to reduce the complexity of the circuits. In 'patch circuits', we remove a slice of two-qubit gates (a small fraction of the total number of two-qubit gates), splitting the circuit into two spatially isolated, non-interacting patches of qubits. We then compute the total fidelity as the product of the patch fidelities, each of which can be easily calculated. In 'elided circuits', we remove only a fraction of the initial two-qubit gates along the slice, allowing for entanglement between patches, which more closely mimics the full experiment while still maintaining simulation feasibility. Finally, we can also run full 'verification circuits', with the same gate counts as our supremacy circuits, but with a different pattern for the sequence of two-qubit gates, which is much easier to simulate classically (see also Supplementary Information). Comparison between these three variations allows us to track the system fidelity as we approach the supremacy regime.

We first check that the patch and elided versions of the verification circuits produce the same fidelity as the full verification circuits up to 53 qubits, as shown in Fig. 4a. For each data point, we typically collect $N_s = 5 \times 10^6$ total samples over ten circuit instances, where instances differ only in the choices of single-qubit gates in each cycle. We also show predicted $\mathcal{F}_{\mathrm{XEB}}$ values, computed by multiplying the no-error probabilities of single- and two-qubit gates and measurement (see also Supplementary Information). The predicted, patch and elided fidelities all show good agreement with the fidelities of the corresponding full circuits, despite the vast differences in computational complexity and entanglement. This gives us confidence that elided circuits can be used to accurately estimate the fidelity of more-complex circuits.

The largest circuits for which the fidelity can still be directly verified have 53 qubits and a simplified gate arrangement. Performing random circuit sampling on these at 0.8% fidelity takes one million cores 130 seconds, corresponding to a million-fold speedup of the quantum processor relative to a single core.

We proceed now to benchmark our computationally most difficult circuits, which are simply a rearrangement of the two-qubit gates. In Fig. 4b, we show the measured $\mathcal{F}_{\mathrm{XEB}}$ for 53-qubit patch and elided versions of the full supremacy circuits with increasing depth. For the largest circuit with 53 qubits and 20 cycles, we collected $N_s = 30 \times 10^6$ samples over ten circuit instances, obtaining $\mathcal{F}_{\mathrm{XEB}} = (2.24 \pm 0.21) \times 10^{-3}$ for the elided circuits. With $5\sigma$ confidence, we assert that the average fidelity of running these circuits on the quantum processor is greater than at least 0.1%. We expect that the full data for Fig. 4b should have similar fidelities, but since the simulation times (red numbers) take too long to check, we have archived the data (see 'Data availability' section). The data is thus in the quantum supremacy regime.

## The classical computational cost

We simulate the quantum circuits used in the experiment on classical computers for two purposes: (1) verifying our quantum processor and benchmarking methods by computing $\mathcal{F}_{\mathrm{XEB}}$ where possible using simplifiable circuits (Fig. 4a), and (2) estimating $\mathcal{F}_{\mathrm{XEB}}$ as well as the classical cost of sampling our hardest circuits (Fig. 4b). Up to 43 qubits, we use a Schrödinger algorithm, which simulates the evolution of the full quantum state; the Jülich supercomputer (with 100,000 cores, 250 terabytes) runs the largest cases. Above this size, there is not enough random access memory (RAM) to store the quantum state[42]. For larger qubit numbers, we use a hybrid Schrödinger–Feynman algorithm[43] running on Google data centres to compute the amplitudes of individual bitstrings. This algorithm breaks the circuit up into two patches of qubits and efficiently simulates each patch using a Schrödinger method, before connecting them using an approach reminiscent of the Feynman path-integral. Although it is more memory-efficient, the Schrödinger–Feynman algorithm becomes exponentially more computationally expensive with increasing circuit depth owing to the exponential growth of paths with the number of gates connecting the patches.

To estimate the classical computational cost of the supremacy circuits (grey numbers in Fig. 4b), we ran portions of the quantum circuit simulation on both the Summit supercomputer as well as on Google clusters and extrapolated to the full cost. In this extrapolation, we account for the computation cost of sampling by scaling the verification cost with $\mathcal{F}_{\mathrm{XEB}}$, for example[43,44], a 0.1% fidelity decreases the cost by about 1,000. On the Summit supercomputer, which is currently the most powerful in the world, we used a method inspired by Feynman path-integrals that is most efficient at low depth[44–47]. At $m = 20$ the tensors do not reasonably fit into node memory, so we can only measure runtimes up to $m = 14$, for which we estimate that sampling three million bitstrings with 1% fidelity would require a year.

On Google Cloud servers, we estimate that performing the same task for $m = 20$ with 0.1% fidelity using the Schrödinger–Feynman algorithm would cost 50 trillion core-hours and consume one petawatt hour of energy. To put this in perspective, it took 600 seconds to sample the circuit on the quantum processor three million times, where sampling time is limited by control hardware communications; in fact, the net

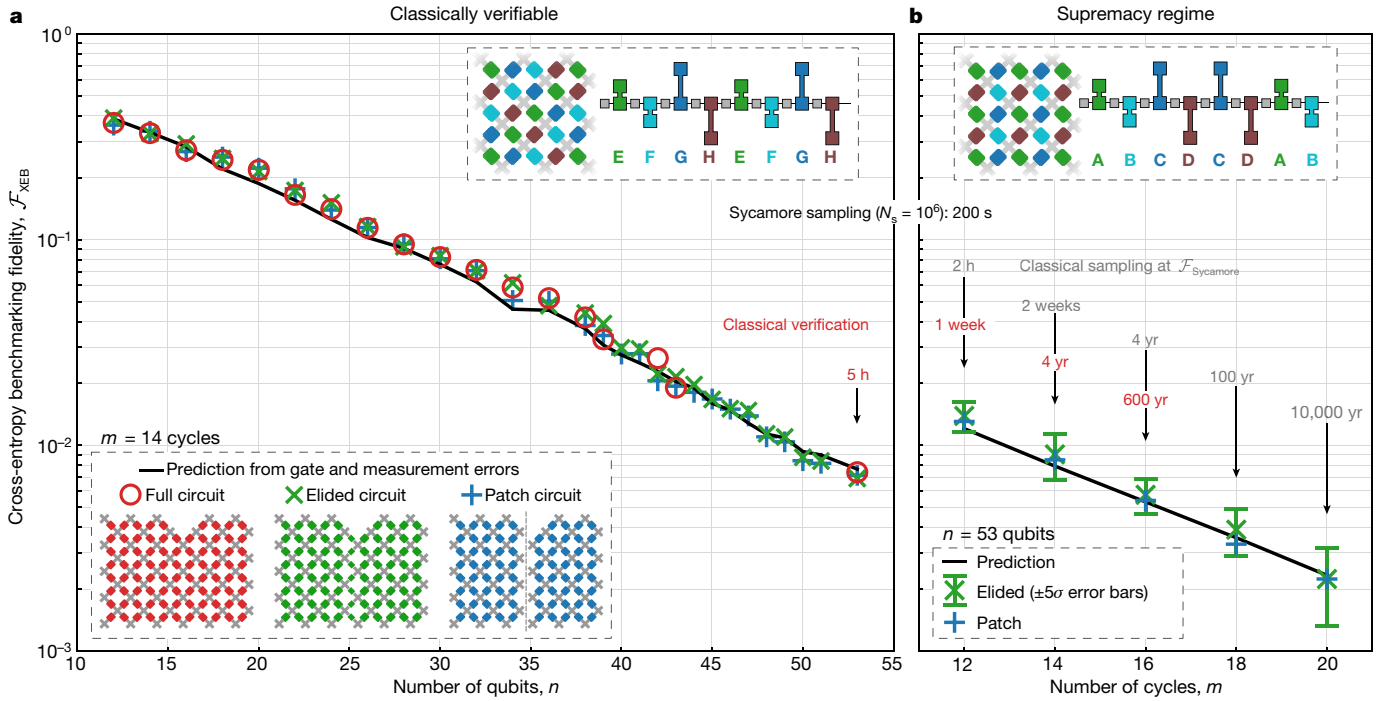



**Fig. 4 | Demonstrating quantum supremacy. a**, Verification of benchmarking methods. $\mathcal{F}_{XEB}$ values for patch, elided and full verification circuits are calculated from measured bitstrings and the corresponding probabilities predicted by classical simulation. Here, the two-qubit gates are applied in a simplifiable tiling and sequence such that the full circuits can be simulated out to $n = 53$, $m = 14$ in a reasonable amount of time. Each data point is an average over ten distinct quantum circuit instances that differ in their single-qubit gates (for $n = 39$, 42 and 43 only two instances were simulated). For each $n$, each instance is sampled with $N_s$ of 0.5–2.5 million. The black line shows the predicted $\mathcal{F}_{XEB}$ based on single- and two-qubit gate and measurement errors. The close correspondence between all four curves, despite their vast differences in complexity, justifies the use of elided circuits to estimate fidelity in the supremacy regime. **b**, Estimating $\mathcal{F}_{XEB}$ in the quantum supremacy regime. Here, the two-qubit gates are applied in a non-simplifiable tiling and sequence for which it is much harder to simulate. For the largest elided data ($n = 53$, $m = 20$, total $N_s = 30$ million), we find an average $\mathcal{F}_{XEB} > 0.1\%$ with $5\sigma$ confidence, where $\sigma$ includes both systematic and statistical uncertainties. The corresponding full circuit data, not simulated but archived, is expected to show similarly statistically significant fidelity. For $m = 20$, obtaining a million samples on the quantum processor takes 200 seconds, whereas an equal-fidelity classical sampling would take 10,000 years on a million cores, and verifying the fidelity would take millions of years.

quantum processor time is only about 30 seconds. The bitstring samples from all circuits have been archived online (see 'Data availability' section) to encourage development and testing of more advanced verification algorithms.

One may wonder to what extent algorithmic innovation can enhance classical simulations. Our assumption, based on insights from complexity theory[11–13], is that the cost of this algorithmic task is exponential in circuit size. Indeed, simulation methods have improved steadily over the past few years[42–50]. We expect that lower simulation costs than reported here will eventually be achieved, but we also expect that they will be consistently outpaced by hardware improvements on larger quantum processors.

## Verifying the digital error model

A key assumption underlying the theory of quantum error correction is that quantum state errors may be considered digitized and localized[38,51]. Under such a digital model, all errors in the evolving quantum state may be characterized by a set of localized Pauli errors (bit-flips or phase-flips) interspersed into the circuit. Since continuous amplitudes are fundamental to quantum mechanics, it needs to be tested whether errors in a quantum system could be treated as discrete and probabilistic. Indeed, our experimental observations support the validity of this model for our processor. Our system fidelity is well predicted by a simple model in which the individually characterized fidelities of each gate are multiplied together (Fig. 4).

To be successfully described by a digitized error model, a system should be low in correlated errors. We achieve this in our experiment by choosing circuits that randomize and decorrelate errors, by optimizing control to minimize systematic errors and leakage, and by designing gates that operate much faster than correlated noise sources, such as $1/f$ flux noise[37]. Demonstrating a predictive uncorrelated error model up to a Hilbert space of size $2^{53}$ shows that we can build a system where quantum resources, such as entanglement, are not prohibitively fragile.

## The future

Quantum processors based on superconducting qubits can now perform computations in a Hilbert space of dimension $2^{53} \approx 9 \times 10^{15}$, beyond the reach of the fastest classical supercomputers available today. To our knowledge, this experiment marks the first computation that can be performed only on a quantum processor. Quantum processors have thus reached the regime of quantum supremacy. We expect that their computational power will continue to grow at a double-exponential rate: the classical cost of simulating a quantum circuit increases exponentially with computational volume, and hardware improvements will probably follow a quantum-processor equivalent of Moore's law[52,53], doubling this computational volume every few years. To sustain the double-exponential growth rate and to eventually offer the computational volume needed to run well known quantum algorithms, such as the Shor or Grover algorithms[25,54], the engineering of quantum error correction will need to become a focus of attention.

The extended Church–Turing thesis formulated by Bernstein and Vazirani[55] asserts that any 'reasonable' model of computation can be efficiently simulated by a Turing machine. Our experiment suggests that a model of computation may now be available that violates this

assertion. We have performed random quantum circuit sampling in polynomial time using a physically realizable quantum processor (with sufficiently low error rates), yet no efficient method is known to exist for classical computing machinery. As a result of these developments, quantum computing is transitioning from a research topic to a technology that unlocks new computational capabilities. We are only one creative algorithm away from valuable near-term applications.

## Data availability

The datasets generated and analysed for this study are available at our public Dryad repository (https://doi.org/10.5061/dryad.k6t1rj8).

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1666-5.


1. Feynman, R. P. Simulating physics with computers. *Int. J. Theor. Phys.* **21**, 467–488 (1982).
2. Devoret, M. H., Martinis, J. M. & Clarke, J. Measurements of macroscopic quantum tunneling out of the zero-voltage state of a current-biased Josephson junction. *Phys. Rev. Lett.* **55**, 1908 (1985).
3. Nakamura, Y., Chen, C. D. & Tsai, J. S. Spectroscopy of energy-level splitting between two macroscopic quantum states of charge coherently superposed by Josephson coupling. *Phys. Rev. Lett.* **79**, 2328 (1997).
4. Mooij, J. et al. Josephson persistent-current qubit. *Science* **285**, 1036–1039 (1999).
5. Wallraff, A. et al. Strong coupling of a single photon to a superconducting qubit using circuit quantum electrodynamics. *Nature* **431**, 162–167 (2004).
6. Koch, J. et al. Charge-insensitive qubit design derived from the Cooper pair box. *Phys. Rev. A* **76**, 042319 (2007).
7. You, J. Q. & Nori, F. Atomic physics and quantum optics using superconducting circuits. *Nature* **474**, 589–597 (2011).
8. Preskill, J. Quantum computing and the entanglement frontier. *Rapporteur Talk at the 25th Solvay Conference on Physics, Brussels* https://doi.org/10.1142/8674 (World Scientific, 2012).
9. Aaronson, S. & Arkhipov, A. The computational complexity of linear optics. In *Proc. 43rd Ann. Symp. on Theory of Computing* https://doi.org/10.1145/1993636.1993682 (ACM, 2011).
10. Bremner, M. J., Montanaro, A. & Shepherd, D. J. Average-case complexity versus approximate simulation of commuting quantum computations. *Phys. Rev. Lett.* **117**, 080501 (2016).
11. Boixo, S. et al. Characterizing quantum supremacy in near-term devices. *Nat. Phys.* **14**, 595 (2018).
12. Bouland, A., Fefferman, B., Nirkhe, C. & Vazirani, U. On the complexity and verification of quantum random circuit sampling. *Nat. Phys.* **15**, 159 (2019).
13. Aaronson, S. & Chen, L. Complexity-theoretic foundations of quantum supremacy experiments. In *32nd Computational Complexity Conf.* https://doi.org/10.4230/LIPIcs.CCC.2017.22 (Schloss Dagstuhl–Leibniz Zentrum für Informatik, 2017).
14. Neill, C. et al. A blueprint for demonstrating quantum supremacy with superconducting qubits. *Science* **360**, 195–199 (2018).
15. Preskill, J. Quantum computing in the NISQ era and beyond. *Quantum* **2**, 79 (2018).
16. Kechedzhi, K. et al. Efficient population transfer via non-ergodic extended states in quantum spin glass. In *13th Conf. on the Theory of Quantum Computation, Communication and Cryptography* http://drops.dagstuhl.de/opus/volltexte/2018/9256/pdf/LIPIcs-TQC-2018-9.pdf (Schloss Dagstuhl–Leibniz Zentrum für Informatik, 2018).
17. Somma, R. D., Boixo, S., Barnum, H. & Knill, E. Quantum simulations of classical annealing processes. *Phys. Rev. Lett.* **101**, 130504 (2008).
18. Farhi, E. & Neven, H. Classification with quantum neural networks on near term processors. Preprint at https://arxiv.org/abs/1802.06002 (2018).
19. McClean, J. R., Boixo, S., Smelyanskiy, V. N., Babbush, R. & Neven, H. Barren plateaus in quantum neural network training landscapes. *Nat. Commun.* **9**, 4812 (2018).
20. Cong, I., Choi, S. & Lukin, M. D. Quantum convolutional neural networks. *Nat. Phys.* https://doi.org/10.1038/s41567-019-0648-8 (2019).
21. Bravyi, S., Gosset, D. & König, R. Quantum advantage with shallow circuits. *Science* **362**, 308–311 (2018).
22. Aspuru-Guzik, A., Dutoi, A. D., Love, P. J. & Head-Gordon, M. Simulated quantum computation of molecular energies. *Science* **309**, 1704–1707 (2005).
23. Peruzzo, A. et al. A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun.* **5**, 4213 (2014).
24. Hempel, C. et al. Quantum chemistry calculations on a trapped-ion quantum simulator. *Phys. Rev. X* **8**, 031022 (2018).
25. Shor, P. W. Algorithms for quantum computation: discrete logarithms and factoring proceedings. In *Proc. 35th Ann. Symp. on Foundations of Computer Science* https://doi.org/10.1109/SFCS.1994.365700 (IEEE, 1994).
26. Fowler, A. G., Mariantoni, M., Martinis, J. M. & Cleland, A. N. Surface codes: towards practical large-scale quantum computation. *Phys. Rev. A* **86**, 032324 (2012).
27. Barends, R. et al. Superconducting quantum circuits at the surface code threshold for fault tolerance. *Nature* **508**, 500–503 (2014).
28. Córcoles, A. D. et al. Demonstration of a quantum error detection code using a square lattice of four superconducting qubits. *Nat. Commun.* **6**, 6979 (2015).
29. Ofek, N. et al. Extending the lifetime of a quantum bit with error correction in superconducting circuits. *Nature* **536**, 441 (2016).
30. Vool, U. & Devoret, M. Introduction to quantum electromagnetic circuits. *Int. J. Circuit Theory Appl.* **45**, 897–934 (2017).
31. Chen, Y. et al. Qubit architecture with high coherence and fast tunable coupling circuits. *Phys. Rev. Lett.* **113**, 220502 (2014).
32. Yan, F. et al. A tunable coupling scheme for implementing high-fidelity two-qubit gates. *Phys. Rev. Appl.* **10**, 054062 (2018).
33. Schuster, D. I. et al. Resolving photon number states in a superconducting circuit. *Nature* **445**, 515 (2007).
34. Jeffrey, E. et al. Fast accurate state measurement with superconducting qubits. *Phys. Rev. Lett.* **112**, 190504 (2014).
35. Chen, Z. et al. Measuring and suppressing quantum state leakage in a superconducting qubit. *Phys. Rev. Lett.* **116**, 020501 (2016).
36. Klimov, P. V. et al. Fluctuations of energy-relaxation times in superconducting qubits. *Phys. Rev. Lett.* **121**, 090502 (2018).
37. Yan, F. et al. The flux qubit revisited to enhance coherence and reproducibility. *Nat. Commun.* **7**, 12964 (2016).
38. Knill, E. et al. Randomized benchmarking of quantum gates. *Phys. Rev. A* **77**, 012307 (2008).
39. Magesan, E., Gambetta, J. M. & Emerson, J. Scalable and robust randomized benchmarking of quantum processes. *Phys. Rev. Lett.* **106**, 180504 (2011).
40. Cross, A. W., Magesan, E., Bishop, L. S., Smolin, J. A. & Gambetta, J. M. Scalable randomised benchmarking of non-Clifford gates. *npj Quant. Inform.* **2**, 16012 (2016).
41. Wallraff, A. et al. Approaching unit visibility for control of a superconducting qubit with dispersive readout. *Phys. Rev. Lett.* **95**, 060501 (2005).
42. De Raedt, H. et al. Massively parallel quantum computer simulator, eleven years later. *Comput. Phys. Commun.* **237**, 47–61 (2019).
43. Markov, I. L., Fatima, A., Isakov, S. V. & Boixo, S. Quantum supremacy is both closer and farther than it appears. Preprint at https://arxiv.org/abs/1807.10749 (2018).
44. Villalonga, B. et al. A flexible high-performance simulator for the verification and benchmarking of quantum circuits implemented on real hardware. *npj Quant. Inform.* (in the press); preprint at https://arxiv.org/abs/1811.09599 (2018).
45. Boixo, S., Isakov, S. V., Smelyanskiy, V. N. & Neven, H. Simulation of low-depth quantum circuits as complex undirected graphical models. Preprint at https://arxiv.org/abs/1712.05384 (2017).
46. Chen, J., Zhang, F., Huang, C., Newman, M. & Shi, Y. Classical simulation of intermediate-size quantum circuits. Preprint at https://arxiv.org/abs/1805.01450 (2018).
47. Villalonga, B. et al. Establishing the quantum supremacy frontier with a 281 pflop/s simulation. Preprint at https://arxiv.org/abs/1905.00444 (2019).
48. Pednault, E. et al. Breaking the 49-qubit barrier in the simulation of quantum circuits. Preprint at https://arxiv.org/abs/1710.05867 (2017).
49. Chen, Z. Y. et al. 64-qubit quantum circuit simulation. *Sci. Bull.* **63**, 964–971 (2018).
50. Chen, M.-C. et al. Quantum-teleportation-inspired algorithm for sampling large random quantum circuits. Preprint at https://arxiv.org/abs/1901.05003 (2019).
51. Shor, P. W. Scheme for reducing decoherence in quantum computer memory. *Phys. Rev. A* **52**, R2493–R2496 (1995).
52. Devoret, M. H. & Schoelkopf, R. J. Superconducting circuits for quantum information: an outlook. *Science* **339**, 1169–1174 (2013).
53. Mohseni, M. et al. Commercialize quantum technologies in five years. *Nature* **543**, 171 (2017).
54. Grover, L. K. Quantum mechanics helps in searching for a needle in a haystack. *Phys. Rev. Lett.* **79**, 325 (1997).
55. Bernstein, E. & Vazirani, U. Quantum complexity theory. In *Proc. 25th Ann. Symp. on Theory of Computing* https://doi.org/10.1145/167088.167097 (ACM, 1993).

**Acknowledgements** We are grateful to E. Schmidt, S. Brin, S. Pichai, J. Dean, J. Yagnik and J. Giannandrea for their executive sponsorship of the Google AI Quantum team, and for their continued engagement and support. We thank P. Norvig, J. Yagnik, U. Hölzle and S. Pichai for advice on the manuscript. We acknowledge K. Kissel, J. Raso, D. L. Yonge-Mallo, O. Martin and N. Sridhar for their help with simulations. We thank G. Bortoli and L. Laws for keeping our team organized. This research used resources from the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility (supported by contract DE-AC05-00OR22725). A portion of this work was performed in the UCSB Nanofabrication Facility, an open access laboratory. R.B., S.M., and E.G.R. appreciate support from the NASA Ames Research Center and from the Air Force Research (AFRL) Information Directorate (grant F4HBKC4162G001). T.S.H. is supported by the DOE Early Career Research Program. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of AFRL or the US government.

**Author contributions** The Google AI Quantum team conceived the experiment. The applications and algorithms team provided the theoretical foundation and the specifics of the algorithm. The hardware team carried out the experiment and collected the data. The data analysis was done jointly with outside collaborators. All authors wrote and revised the manuscript and the Supplementary Information.

**Competing interests** The authors declare no competing interests.

**Additional information**

**Supplementary information** is available for this paper at https://doi.org/10.1038/s41586-019-1666-5.

**Correspondence and requests for materials** should be addressed to J.M.M.

**Peer review information** *Nature* thanks Scott Aaronson, Keisuke Fujii and William Oliver for their contribution to the peer review of this work.

Article

# An antiaromatic-walled nanospace

https://doi.org/10.1038/s41586-019-1661-x

Received: 1 April 2019

Accepted: 14 August 2019

Published online: 23 October 2019

Masahiro Yamashina[1,4], Yuya Tanaka[2], Roy Lavendomme[1], Tanya K. Ronson[1], Michael Pittelkow[3] & Jonathan R. Nitschke[1]*

Over the past few decades, several molecular cages, hosts and nanoporous materials enclosing nanometre-sized cavities have been reported[1–5], including coordination-driven nanocages[6]. Such nanocages have found widespread use in molecular recognition, separation, stabilization and the promotion of unusual chemical reactions, among other applications[3–10]. Most of the reported nanospaces within molecular hosts are confined by aromatic walls, the properties of which help to determine the host–guest behaviour. However, cages with nanospaces surrounded by antiaromatic walls have not yet been developed, owing to the instability of antiaromatic compounds; as such, the effect of antiaromatic walls on the properties of nanospaces remains unknown. Here we demonstrate the construction of an antiaromatic-walled nanospace within a self-assembled cage composed of four metal ions with six identical antiaromatic walls. Calculations indicate that the magnetic effects of the antiaromatic moieties surrounding this nanospace reinforce each other. This prediction is confirmed by $^{1}$H nuclear magnetic resonance (NMR) signals of bound guest molecules, which are observed at chemical shift values of up to 24 parts per million (ppm), owing to the combined antiaromatic deshielding effect of the surrounding rings. This value, shifted 15 ppm from that of the free guest, is the largest $^{1}$H NMR chemical shift displacement resulting from an antiaromatic environment observed so far. This cage may thus be considered as a type of NMR shift reagent, moving guest signals well beyond the usual NMR frequency range and opening the way to further probing the effects of an antiaromatic environment on a nanospace.

Aromaticity and antiaromaticity are fundamental concepts in chemistry, and a long-standing challenge is the preparation of antiaromatic molecules and the study of their properties. Most cavities within coordination cages can be considered as 'aromatic-walled nanospaces' owing to the aromatic character of the surrounding walls[3–5,10] (Fig. 1a). Aromatic-walled nanospaces are characterized by an intermolecular aromatic NMR-shielding effect, in which the nuclei of included guests experience a weaker magnetic field than the one applied[11], with endohedral fullerenes[12] providing a canonical example of such NMR effects. By contrast, a cavity surrounded by antiaromatic walls would experience deshielding (Fig. 1b), or enhancement of the external magnetic field, because an antiaromatic ring generates an induced magnetic field in the opposite direction to that of an aromatic ring. The creation of an antiaromatic-walled nanospace is a challenging task because it requires the precise placement of unstable antiaromatic walls around the central cavity. The instability of these walls, reflected in a high degree of chemical reactivity, is a consequence of the electronic structure of antiaromatic molecules[13,14].

Antiaromatic compounds, which have a cyclic and planar π-conjugated system with $4n$ π electrons, are generally quite reactive: electrons are relatively easy both to add and to remove, and such compounds may react as diradicals[15]. To utilize antiaromatic rings as building blocks for antiaromatic-walled nanospaces, both high stability and strong antiaromaticity are required—a set of properties that initially seem mutually incompatible. After investigating different preparation strategies that have been reported recently for antiaromatic systems[16–20], we came upon the work of Shinokubo and co-workers[21], who reported the facile synthesis of an antiaromatic porphyrinoid with 16 π electrons, $Ni^{II}$-dimesitylnorcorrole (**1**; Extended Data Fig. 1). Despite the strong antiaromaticity over the entire surface of the central moiety of **1**, it is stable under ambient conditions. This unusual stability permitted access to functionalization of **1** by several chemical reactions[22–26]. We thus set out to use **1** as a building block for an antiaromatic-walled nanospace through subcomponent self-assembly[27].

Di(aniline)-based subcomponent **2** was synthesized from **1** in three steps on the basis of published methods[22,26] (Extended Data Fig. 1). Antiaromatic cage **3** was then constructed using subcomponent self-assembly. Diamine **2** (6 equiv.), 2-formylpyridine (12 equiv.) and $Fe^{II}$ bis(trifluoromethanesulfonyl)imide ($NTf_2^-$) (4 equiv.) were mixed in $CH_3CN$, resulting in the formation of $Fe^{II}_4L_6$ cage **3** as the uniquely observed product (Fig. 2a). The structure of **3** was characterized by NMR spectroscopy, mass spectrometry (MS) and X-ray crystallography. All of the $^{1}$H NMR signals were assigned using different NMR techniques (Supplementary Figs. 10–21). A set of signals for norcorrole moieties ($H_{d-f}$) was observed at 1.76–2.02 ppm, in the same high-field region as **1** and **2**, indicating that the antiaromaticity of the norcorrole skeleton was retained following assembly into cage **3** (Fig. 2c, Supplementary

[1]Department of Chemistry, University of Cambridge, Cambridge, UK. [2]Laboratory for Chemistry and Life Science, Institute of Innovative Research, Tokyo Institute of Technology, Yokohama, Japan. [3]Department of Chemistry, University of Copenhagen, Copenhagen, Denmark. [4]Present address: Department of Chemistry, School of Science, Tokyo Institute of Technology, Tokyo, Japan. *e-mail: jrn34@cam.ac.uk

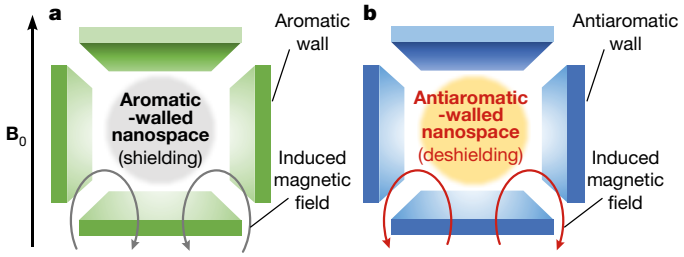



**Fig. 1 | Cartoon representations of nanospaces. a**, **b**, An aromatic-walled nanospace surrounded by aromatic walls (**a**) and an antiaromatic-walled nanospace with antiaromatic walls (**b**). $\mathbf{B}_0$ is the applied magnetic field.

Fig. 10). The signals of the bulky mesityl groups ($H_{b,c}$), which cannot freely rotate, split in two, corresponding to environments inside and outside the cage. The phenylene signals ($H_{g,h}$) were considerably broadened at room temperature. We thus infer that phenylene rotation is restricted by the nearby mesityl and pyridine moieties. However, a set of four sharp phenylene doublets was observed at 243 K, which correspond to protons inside and outside the cage, as with the mesityl signals (Fig. 2d, Supplementary Fig. 11). The diffusion-ordered spectroscopy (DOSY) NMR spectrum shows a single band with a diffusion coefficient ($D$) of $3.98 \times 10^{-10}$ m$^2$ s$^{-1}$, which corresponds to a diameter of **3** of about 3 nm (Supplementary Fig. 21). Prominent signals of the antiaromatic cage were observed under standard electrospray ionization time-of-flight (ESI-TOF) MS conditions (Supplementary Fig. 22).

X-ray crystallographic analysis provided unambiguous evidence for the formation of the $Fe^{II}_4L_6$ structure of **3**. Dark-red single crystals of **3** were obtained by slow diffusion of $Et_2O$ into a solution of **3** in $CH_3CN$. Six ligands bridge four octahedral $Fe^{II}$ centres to provide the expected tetrahedral cage with $T$ symmetry (Fig. 3a, Supplementary Figs. 23, 24, Supplementary Table 1), with four apertures of ~3.3 Å (between proximal $H_d$–$H_e$ on adjacent norcorrole units) on the faces. All $Fe^{II}$ centres in each tetrahedron have the same handedness (ΔΔΔΔ or ΛΛΛΛ). The metal–metal distances are 21.9 Å for Fe···Fe and 14.6 Å between Ni···Ni antipodes. Each norcorrole wall displays a 165.4(3)° (uncertainties are 1$\sigma$) bend inwards (measured as the C3–Ni–C12 angles; Supplementary Fig. 25). Although norcorrole **1** and previously reported 3,12-substituted norcorroles[24] are planar (180.0°), the ditopic norcorrole walls of **3** bow inwards. As a result of this bending, the face apertures are minimized via stacking between mesityl groups and the neighbouring norcorrole edges in the crystal. This conformation is also present in solution, as indicated by nuclear Overhauser effect spectroscopy (NOESY), where correlations are observed between the mesityl and norcorrole moieties ($H^{in}_c$–$H_{d,f}$ and $H^{in}_b$–$H_{d,f}$) of **3** (Supplementary Figs. 17, 26). The cavity volume for the X-ray crystal structure was estimated using the PLATON program to be 1,150 Å$^3$ (Supplementary Fig. 24), substantially less than the 1,950 Å$^3$

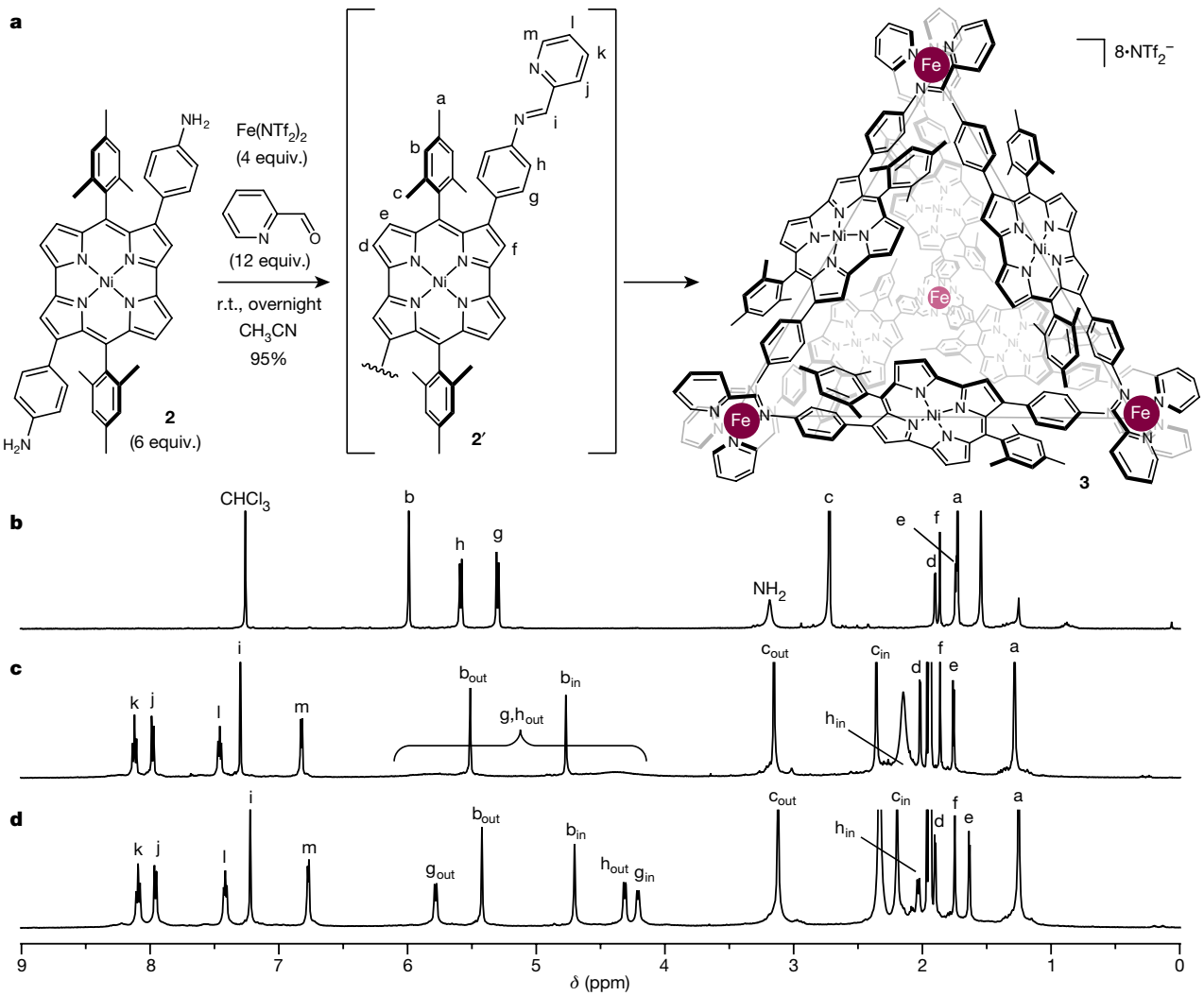



**Fig. 2 | Synthesis and NMR characterization of 3. a**, Subcomponent self-assembly of antiaromatic cage **3**. **b**–**d**, $^1$H NMR spectra (500 MHz) of subcomponent **2** in $CDCl_3$ at 298 K (**b**) and **3** in $CD_3CN$ at 298 K (**c**) and 243 K (**d**). r.t., room temperature.

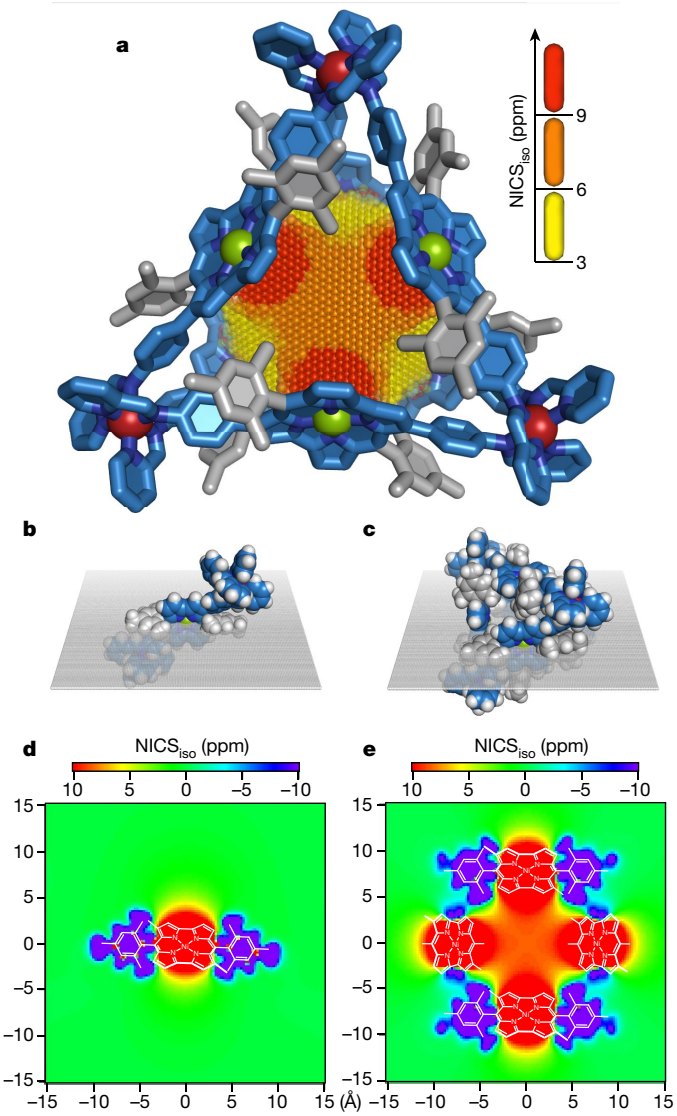



**Fig. 3 | Crystal structure of 3 and NICS calculations. a**, X-ray crystal structure of **3** (ΔΔΔΔ enantiomer) in stick representation with a three-dimensional NICS grid, showing the magnetic deshielding experienced within the antiaromatic-walled nanospace (displayed only for NICS$_{iso}$ > 3; a front layer was sliced off to show the inside region. A more complete view of this three-dimensional NICS grid is shown in the Supplementary Video). **b**, **c**, Cross-sections of bis(Fe$^{II}$) model complex **3′** (**b**) and **3** (**c**), obtained from two-dimensional NICS calculations. **d**, **e**, NICS slice plots of **3′** (**d**) and **3** (**e**) on a 0.25-Å-resolution grid, calculated using the crystal structure at the B3LYP/SDD (for Ni, Fe) and 6-31G(d) (for C, N, H) levels. Red and blue represent deshielding (positive) and shielding (negative) zones, respectively. The chemical structure is overlaid in white as a visual aid.

volume of a model having planar norcorrole walls (Supplementary Fig. 27).

The ultraviolet–visible–near-infrared (UV–vis–NIR) spectra of **3** and **2** are shown in Supplementary Fig. 28. Subcomponent **2** displayed broad absorption bands around 600 nm and 1,000 nm. The former is assigned to intramolecular charge transfer from HOMO − 4 to the LUMO[26], and the latter is a characteristic band for antiaromatic porphyrinoids[21]. These peaks were substantially broadened following cage formation in the spectrum of **3**. These results are consistent with time-dependent density functional theory (TD-DFT) calculations of the spectra of **2** and **3′**. Complex **3′** is a computable model of a single edge of **3**: a bis(Fe$^{II}$) complex, in which each iron centre is bound to one edge of a **2′** ligand and two (*E*)-*N*-phenyl-(pyridin-2-yl)methanimine.

We carried out cyclic voltammetry experiments (Supplementary Fig. 29) to investigate the electrochemical properties of **2** and **3**. Whereas subcomponent **2** displayed three reversible reduction peaks and one irreversible oxidation peak, attributed to oxidation of the primary amines, cage **3** gave a more complex collection of oxidation and reduction waves. The reduction peak at −2.29 V was attributed to the Fe$^{II}$-coordinated pyridylimines[28]. The multiple overlapping waves from −2.0 to +0.5 V probably arose from electrochemical communication among the norcorrole walls of **3**. The calculated HOMO–LUMO gaps of **1**, **2** and **3′** are the same (that is, 1.5 eV; Supplementary Fig. 30), which is consistent with **1** and **2** having the same gap between their first reduction and oxidation (Supplementary Fig. 29).

To investigate the extent of the antiaromaticity experienced within the void volume of **3**, we carried out nucleus-independent chemical shift (NICS) calculations. Calculated NICS(0) values of the norcorrole moieties are given in Supplementary Fig. 31. For the norcorrole walls in **1**, **2** and **3**, the large positive values at the centre of each ring that includes the nickel atom are consistent with strong cavity-wall antiaromaticity. This result prompted us to seek further insight into the environment of the central nanospace. The two-dimensional NICS$_{iso}$ plot of **3** orthogonal to the ring revealed an enhanced antiaromaticity-induced magnetic field within the cavity in comparison with model complex **3′** (Fig. 3b–e). The graphic (Fig. 3a) and animation (Supplementary Video) provide a three-dimensional visualization of this field, as calculated using a three-dimensional NICS$_{iso}$ grid. The calculated NICS(*n*) at the centroid of **3** is $\delta$ = +7.4 ppm, which is approximately six times larger than the NICS(*n*) value of a corresponding point at the same distance from **3′** (Supplementary Fig. 32a, b). This result indicates that the six norcorrole walls of **3** have an additive effect on the antiaromaticity influence of the central environment. Furthermore, positive NICS(*n*) values around the centroid within **3** were consistently high, whereas those at the same coordinates of **3′** decreased as the distance increased from the Ni centre (Supplementary Fig. 32c). These computational studies thus support cooperativity between the antiaromatic walls in increasing the antiaromatic character within **3**. The anisotropy of the induced current density (ACID) of **1**, **2** and **3**, which traces out ring currents, also supports the results of NICS (Supplementary Fig. 33).

Host–guest studies were conducted to investigate experimentally the effect of guest binding within the antiaromatic-walled cavity of **3**. When **3** (1.0 mg, 0.12 μmol) and coronene (**4**; 5 equiv.) were mixed in $CD_3CN$, formation of the 1:2 host–guest complex **3**·(**4**)$_2$ was observed by MS (Supplementary Fig. 37) and NMR (Fig. 4a) analyses. In the $^1$H NMR spectrum of **3**·(**4**)$_2$, the host signals were observed at similar chemical shift values to those of the empty host. Remarkably, however, the encapsulated guest signal was shifted downfield by 8.1 ppm compared to the free guest (Fig. 4c, Supplementary Fig. 37), as a result of the antiaromatic deshielding effect from the surrounding norcorrole walls. This downfield signal was observed to diffuse at the same rate as **3** in the DOSY spectrum (Supplementary Fig. 36). A heteronuclear single quantum coherence (HSQC) spectrum allowed this new signal to be unambiguously assigned to the encapsulated coronenes (Supplementary Fig. 35).

Similarly, the treatment of corannulene (**5**), dibenzo(*g*,*p*)chrysene (**6**), truxene (**7**), carbon nanobelt (**8**)[29] and *N*-methylfulleropyrrolidine (**9**) with **3** in acetonitrile gave rise to 1:2 or 1:1 host–guest complexes, probably driven by a combination of dispersion forces, aromatic stacking, CH–π interactions and solvophobic effects (Fig. 4b, Supplementary Figs. 38–60). The polycyclic aromatic hydrocarbons **6**–**9**, which have low solubility in acetonitrile, all showed encapsulation after mixing a solution of **3** with an excess of solid guest at room temperature. The signals of all encapsulated guest molecules showed substantial downfield shifting within the antiaromatic-walled nanospace. The two encapsulated molecules of $D_2$-symmetric **6** displayed eight peaks in the range 9–22 ppm ($\Delta\delta$ = +1.7–13.4 ppm; Fig. 4d, Supplementary Figs. 42–46). The signals of the aromatic and aliphatic moieties in $C_{3h}$-symmetric **7** appeared in the downfield region at 8–19 ppm ($\Delta\delta$ = +0.7–13.6 ppm; Fig. 4e,

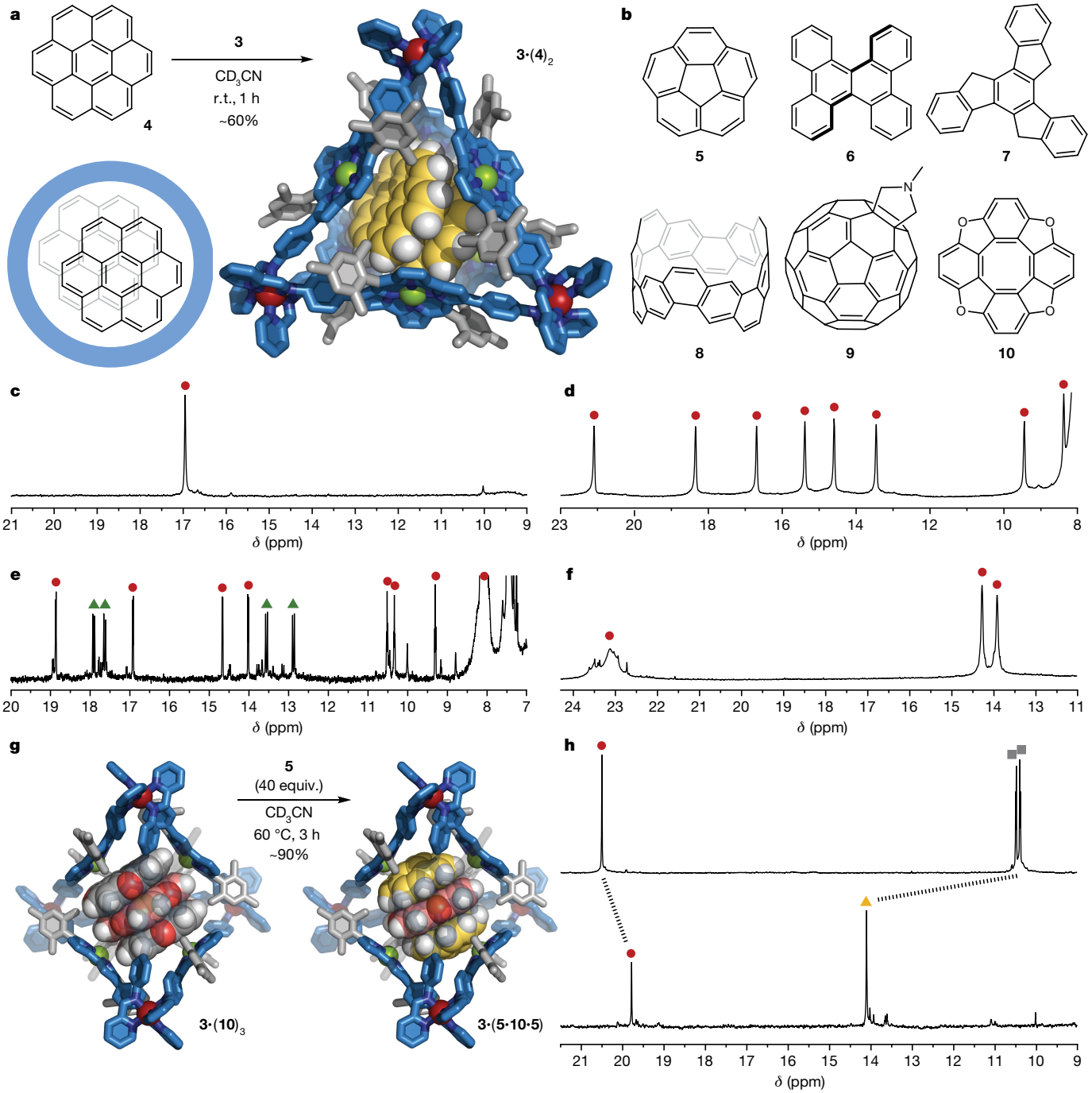



**Fig. 4 | Host–guest chemistry within antiaromatic-walled nanospace. a**, Encapsulation of **4** and an MM3-optimized structure of **3**·(**4**)$_2$ based on the crystal structure of **3**. **b**, Molecules observed to bind within **3**. **c**–**f**, Partial $^1$H NMR spectra (500 MHz, $CD_3CN$, 298 K) of **3**·(**4**)$_2$ (**c**), **3**·(**6**)$_2$ (**d**), **3**·(**7**)$_2$ (**e**) and **3**·**8** (**f**), showing the signals for the encapsulated guests in the downfield region (red circles, aromatic signals; green triangles, aliphatic signals). **g**, Co-encapsulation of **5** and **10**, starting from **3**·(**10**)$_3$ with MM3-optimized structures (the front norcorrole walls are transparent for clarity). **h**, $^1$H NMR spectra (500 MHz, $CD_3CN$, 298 K) of **3**·(**10**)$_3$ (top) and **3**·(**5**·**10**·**5**) (bottom). Red circle, central **10**; grey square, outer **10**; yellow triangle, outer **5**.

Supplementary Figs. 47–52). Two sets of guest signals were observed in the case of **7**. These signals are attributed to diastereomeric dimeric guest configurations within the chiral cavity of **3**. The most extreme downfield shifting was observed for a carbon nanobelt (**8**). The signals of the bay hydrogen atoms appeared at 23 ppm, shifted by +14.9 ppm compared with free **8**, whereas the outer hydrogen signals were shifted by only +6.5 ppm (Fig. 4f, Supplementary Figs. 53–55). Among the set of guests encapsulated (Fig. 4b), the extent of the shift ($\Delta\delta$) varies from 0.7 to 14.9 ppm (Supplementary Fig. 71), depending on the guest and proton position. To obtain further insight into the antiaromaticity-influenced environment within **3**, we conducted NICS calculations at various key points in this nanospace (Supplementary Fig. 32a, b). In contrast to the positive NICS values around the norcorrole walls, the vertex and aperture sites show negative values, which indicate shielding caused by aromaticity. Consequently, protons that are localized near vertices and apertures show relatively minor downfield shifting (Supplementary Figs. 70, 72). The overall association constant $\beta_2$ for the 1:2 host–guest complex **3**·(**5**)$_2$ was estimated to be $4.3 \times 10^5\ M^{-2}$ (±45%) in $CD_3CN$ at 263 K (Supplementary Figs. 73, 74). The association constants for the other guests could not be determined because of their low solubility in acetonitrile.

When **3** and tetraoxa[8]circulene (**10**)[30] were mixed in acetonitrile, the 1:3 host–guest complex **3**·(**10**)$_3$ was formed, without notable effect on guest relaxation times (Fig. 4g, h, Supplementary Figs. 61–67), resulting in full occupation of the antiaromatic-walled nanospace. Within the three-guest stack, the central and outer molecules of **10** are influenced differently by the deshielding effects of the surrounding norcorroles. Whereas the central guest displayed a $\Delta\delta$ shift of +12.8 ppm, the outer ones only shifted by +2.8 ppm.

A remarkable hetero-guest 'sandwich', **3**·(**5**·**10**·**5**), with one circulene intercalated between two corannulenes, was observed to form selectively when **5** (40 equiv.) was added to an acetonitrile solution of **3**·(**10**)$_3$, as confirmed by NMR and electrospray ionization mass spectrometry (ESI-MS) data (Fig. 4g, h, Supplementary Figs. 68–69). Interestingly, the outer corannulenes show large downfield shifts ($\Delta\delta$ = +6.3 ppm), similar to **3**·(**5**)$_2$ ($\Delta\delta$ = +6.6 ppm), in contrast to the smaller downfield shifts of the outer molecules of **10** in **3**·(**10**)$_3$ ($\Delta\delta$ = +2.8 ppm). The central **10** experienced a similar degree of antiaromatic ring current as **3**·(**10**)$_3$ ($\Delta\delta$ = +12.0 ppm versus +12.8 ppm, respectively). We infer that the smaller, concave nature of **5** led to a positioning of its protons within **3**·(**5**·**10**·**5**), which induced a stronger antiaromatic deshielding effect compared to the outer equivalents of **10** in **3**·(**10**)$_3$ (Fig. 4h).

Cage **3** thus serves as a kind of NMR shift reagent that acts without notable shortening of the nuclear relaxation times of guest species, as opposed to paramagnetic NMR shift reagents. This result confirms theoretical predictions of intermolecular effects involving multiple antiaromatic molecules. Future work will explore chemical reactivity within this kind of nanospace.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1661-x.


1. Corma, A. State of the art and future challenges of zeolites as catalysts. *J. Catal.* **216**, 298–312 (2003).
2. Kitagawa, S., Kitaura, R. & Noro, S. Functional porous coordination polymers. *Angew. Chem. Int. Ed.* **43**, 2334–2375 (2004).
3. Hooley, R. J. & Rebek, J. Jr Chemistry and catalysis in functional cavitands. *Chem. Biol.* **16**, 255–264 (2009).
4. Zhang, G. & Mastalerz, M. Organic cage compounds – from shape-persistency to function. *Chem. Soc. Rev.* **43**, 1934–1947 (2014).
5. Cook, T. R., Zheng, Y.-R. & Stang, P. J. Metal-organic frameworks and self-assembled supramolecular coordination complexes: comparing and contrasting the design, synthesis, and functionality of metal-organic materials. *Chem. Rev.* **113**, 734–777 (2013).
6. Fujita, M. et al. Self-assembly of ten molecules into nanometre-sized organic host frameworks. *Nature* **378**, 469–471 (1995).
7. Ariga, K., Ito, H., Hill, J. P. & Tsukube, H. Molecular recognition: from solution science to nano/materials technology. *Chem. Soc. Rev.* **41**, 5800–5835 (2012).
8. Galan, A. & Ballester, P. Stabilization of reactive species by supramolecular encapsulation. *Chem. Soc. Rev.* **45**, 1720–1737 (2016).
9. Yoshizawa, M., Klosterman, J. K. & Fujita, M. Functional molecular flasks: new properties and reactions within discrete, self-assembled hosts. *Angew. Chem. Int. Ed.* **48**, 3418–3438 (2009).
10. Yoshizawa, M. & Yamashina, M. Coordination-driven nanostructures with polyaromatic shells. *Chem. Lett.* **46**, 163–171 (2017).
11. Mugridge, J. S., Bergman, R. G. & Raymond, K. N. $^1$H NMR chemical shift calculations as a probe of supramolecular host–guest geometry. *J. Am. Chem. Soc.* **133**, 11205–11212 (2011).
12. Garcia-Borràs, M., Osuna, S., Luis, J. M., Swart, M. & Solà, M. The role of aromaticity in determining the molecular structure and reactivity of (endohedral metallo)fullerenes. *Chem. Soc. Rev.* **43**, 5089–5105 (2014).
13. Nakamura, Y. et al. A directly fused tetrameric porphyrin sheet and its anomalous electronic properties that arise from the planar cyclooctatetraene core. *J. Am. Chem. Soc.* **128**, 4119–4127 (2006).
14. Peeks, M. D., Claridge, T. D. W. & Anderson, H. L. Aromatic and antiaromatic ring currents in a molecular nanoring. *Nature* **541**, 200–203 (2017).
15. Breslow, R. Antiaromaticity. *Acc. Chem. Res.* **6**, 393–398 (1973).
16. Nishinaga, T., Ohmae, T. & Iyoda, M. Recent studies on the aromaticity and antiaromaticity of planar cyclooctatetraene. *Symmetry (Basel)* **2**, 76–97 (2010).
17. Zeng, Z. et al. Pro-aromatic and anti-aromatic π-conjugated molecules: an irresistible wish to be diradicals. *Chem. Soc. Rev.* **44**, 6578–6596 (2015).
18. Hensel, T., Anderson, N. N., Plesner, M. & Pittelkow, M. Synthesis of heterocyclic [8] circulenes and related structures. *Synlett* **27**, 498–525 (2016).
19. Reddy, B. K., Basavarajappa, A., Ambhore, M. D. & Anand, V. G. Isophlorinoids: the antiaromatic congeners of porphyrinoids. *Chem. Rev.* **117**, 3420–3443 (2017).
20. Szyszko, B., Białek, M. J., Pacholska-Dudziak, E. & Latos-Grażyński, L. Flexible porphyrinoids. *Chem. Rev.* **117**, 2839–2909 (2017).
21. Ito, T., Hayashi, Y., Shimizu, S., Shin, J.-Y., Kobayashi, N. & Shinokubo, H. Gram-scale synthesis of nickel(II) norcorrole: the smallest antiaromatic porphyrinoid. *Angew. Chem. Int. Ed.* **51**, 8542–8545 (2012).
22. Deng, Z., Li, X., Stępień, M. & Chmielewski, P. J. Nitration of norcorrolatonickel(II): first observation of a diatropic current in a system comprising a norcorrole ring. *Chem. Eur. J.* **22**, 4231–4246 (2016).
23. Nozawa, R. et al. Stacked antiaromatic porphyrins. *Nat. Commun.* **7**, 13620 (2016).
24. Yoshida, T. & Shinokubo, H. Direct amination of antiaromatic Ni$^{II}$ norcorrole. *Mater. Chem. Front.* **1**, 1853–1857 (2017).
25. Li, X., Meng, Y., Yi, P., Stępień, M. & Chmielewski, P. J. Pyridine-fused bis(norcorrole) through Hantzsch-type cyclization: enhancement of antiaromaticity by an aromatic bridge. *Angew. Chem. Int. Ed.* **56**, 10810–10814 (2017).
26. Kawashima, H., Hiroto, S. & Shinokubo, H. Acid-mediated migration of bromide in an antiaromatic porphyrinoid: preparation of two regioisomeric Ni(II) bromonorcorroles. *J. Org. Chem.* **82**, 10425–10432 (2017).
27. Zhang, D., Ronson, T. K. & Nitschke, J. R. Functional capsules via subcomponent self-assembly. *Acc. Chem. Res.* **51**, 2423–2436 (2018).
28. Rizzuto, F. J., Wood, D. M., Ronson, T. K. & Nitschke, J. R. Tuning the redox properties of fullerene clusters within a metal–organic capsule. *J. Am. Chem. Soc.* **139**, 11008–11011 (2017).
29. Povie, G., Segawa, Y., Nishihara, T., Miyauchi, Y. & Itami, K. Synthesis of a carbon nanobelt. *Science* **356**, 172–175 (2017).
30. Brock-Nannestad, T. et al. Tetra-*tert*-butyltetraoxa[8]circulene and its unusual aggregation behaviour. *Eur. J. Org. Chem.*, 6320–6325 (2011).

# Methods

### Synthesis of 3,12-di(4-aminophenyl) $Ni^{II}$-dimesitylnorcorrole (2)

To a 100-ml two-necked glass flask, a mixture of di(4-nitrophenyl) $Ni^{II}$-dimesitylnorcorrole **1b** (45.5 mg, 0.0555 mmol), $SnCl_2{\cdot}2H_2O$ (1.25 g, 5.55 mmol) and dry THF (40 ml) was added under $N_2$. The resulting solution was stirred at 70 °C overnight. The mixture was poured in AcOEt (100 ml) and washed with saturated $NaHCO_3$ (aq.), water and brine. The organic phase was dried over $MgSO_4$, filtrated and concentrated under reduced pressure. The crude product was purified by preparative TLC on a silica gel plate with $CH_2Cl_2$ to afford **2** as a dark brown solid (19.4 mg, 46%).

$^1$H NMR (500 MHz, $CDCl_3$, 298 K): $\delta$ 1.73 (s, 6H, *para*-Mes), 1.74 (d, $J$ = 4.0 Hz, 2H, $\beta$-CH), 1.86 (s, 2H, $\beta$-CH), 1.90 (d, $J$ = 4.0 Hz, 2H, $\beta$-CH), 2.72 (s, 12H, *ortho*-Mes), 3.19 (br, 4H, $NH_2$), 5.30 (d, $J$ = 8.5 Hz, 4H, Ph), 5.59 (d, $J$ = 8.5 Hz, 4H, Ph), 5.99 (s, 4H, Mes).

$^{13}$C NMR (125 MHz, $CDCl_3$, 298 K): $\delta$ 18.0 ($CH_3$), 20.6 ($CH_3$), 112.3 (CH), 113.5 (CH), 115.9 (CH), 122.2 ($C_q$), 125.5 (CH), 125.8 (CH), 126.7 ($C_q$), 127.7 (CH), 133.8 ($C_q$), 136.7 ($C_q$), 144.4 ($C_q$), 145.3 ($C_q$), 147.7 ($C_q$), 153.6 ($C_q$), 156.2 ($C_q$), 160.3 ($C_q$), 171.1 ($C_q$).

HR MS (ESI-TOF, $CH_2Cl_2$) *m/z*: $[M]^+$ calcd for $C_{48}H_{40}N_6Ni$, 758.2662; found 758.2648.

### Formation of antiaromatic-walled nanospace 3

Subcomponent **2** (50.0 mg, 0.0658 mmol), $Fe(NTf_2)_2{\cdot}4.5H_2O$ (30.4 mg, 0.0436 mmol), 2-formylpyridine (14.0 mg, 0.131 mmol) and $CH_3CN$ (25 ml) were added to a 50-ml round-bottom flask, and the mixture was stirred at room temperature overnight. The dark-red solution was concentrated to 5 ml under reduced pressure and poured in $Et_2O$ (50 ml). The resulting solid was collected by centrifugation, washed with additional $Et_2O$ and then dried to give **3** as a dark-red solid (83.7 mg, 0.0103 mmol, 95%).

$^1$H NMR (500 MHz, $CD_3CN$, 298 K): $\delta$ 1.28 (s, 36H, *para*-Mes), 1.76 (d, $J$ = 4.0 Hz, 12H, $\beta$-CH), 1.86 (s, 12H, $\beta$-CH), 2.02 (d, $J$ = 4.0 Hz, 12H, $\beta$-CH), 2.15 (br, 12H, Ph), 2.36 (s, 36H, *ortho*-Mes), 3.15 (s, 36H, *ortho*-Mes), 3.78–5.00 (br, 24H, Ph), 4.77 (s, 12H, Mes), 5.27–6.43 (br, 12H, Ph), 5.51 (s, 12H, Mes), 6.82 (d, $J$ = 5.5 Hz, 12H, py), 7.30 (s, 12H, imine), 7.46 (dd, $J$ = 6.5, 5.5 Hz, 12H, py), 7.98 (d, $J$ = 7.5 Hz, 12H, py), 8.12 (dd, $J$ = 7.5, 6.5 Hz, 12H, py).

$^1$H NMR (500 MHz, $CD_3CN$, 243 K): $\delta$ 1.28 (s, 36H, *para*-Mes), 1.67 (d, $J$ = 4.0 Hz, 12H, $\beta$-CH), 1.78 (s, 12H, $\beta$-CH), 1.93 (d, $J$ = 4.0 Hz, 12H, $\beta$-CH), 2.06 (d, $J$ = 8.0 Hz, 12H, Ph), 2.23 (s, 36H, *ortho*-Mes), 3.15 (s, 36H, *ortho*-Mes), 4.24 (d, $J$ = 8.0 Hz, 12H, Ph), 4.34 (d, $J$ = 8.0 Hz, 12H, Ph), 4.73 (s, 12H, Mes), 5.45 (s, 12H, Mes), 5.81 (d, $J$ = 8.0 Hz, 12H, Ph), 6.80 (d, $J$ = 5.5 Hz, 12H, py), 7.25 (s, 12H, imine), 7.45 (dd, $J$ = 6.5, 5.5 Hz, 12H, py), 7.99 (d, $J$ = 7.5 Hz, 12H, py), 8.13 (dd, $J$ = 7.5, 6.5 Hz, 12H, py).

$^{13}$C NMR (125 MHz, $CDCl_3$, 298 K): $\delta$ 17.2 ($CH_3$), 18.8 ($CH_3$), 20.2 ($CH_3$), 116.9 (CH), 116.9 (CH), 117.1 (CH), 119.7 (CH), 122.2 ($C_q$), 124.8 (CH), 126.5 (CH), 127.9 (CH), 128.2 (CH), 130.3 (CH), 131.7 (CH), 132.0 (CH), 132.7 ($C_q$), 133.7 ($C_q$), 134.8 ($C_q$), 137.5 ($C_q$), 140.4 (CH), 146.2 ($C_q$), 147.7 ($C_q$), 148.5 ($C_q$), 149.6 ($C_q$), 156.2 (CH), 158.1 ($C_q$), 159.5 ($C_q$), 168.3 ($C_q$), 169.7 ($C_q$), 173.4 (CH).

$^{13}$C NMR (125 MHz, $CDCl_3$, 243 K): $\delta$ 16.7 ($CH_3$), 18.5 ($CH_3$), 19.8 ($CH_3$), 116.4 (CH), 116.6 (CH), 117.0 (CH), 119.1 (CH), 121.7 ($C_q$), 123.4 (CH), 125.9 (CH), 127.5 (CH), 127.6 (CH), 129.8 (CH), 131.4 (CH), 131.6 (CH), 131.9 ($C_q$), 133.0 ($C_q$), 134.4 ($C_q$), 137.1 ($C_q$), 139.9 (CH), 145.9 ($C_q$), 147.3 ($C_q$), 147.9 ($C_q$), 149.2 ($C_q$), 155.8 (CH), 157.5 ($C_q$), 159.3 ($C_q$), 168.1 ($C_q$), 169.6 ($C_q$), 173.2 (CH).

$^{19}$F NMR (471 MHz, $CD_3CN$, 298 K): $\delta$ −80.5 (s, $CF_3$).

ESI-TOF MS ($CH_3CN$): *m/z* 1742.3 [**3** − $4NTf_2^-$]$^{4+}$, 1337.9 [**3** − $5NTf_2^-$]$^{5+}$, 1068.3 [**3** − $6NTf_2^-$]$^{6+}$, 875.7 [**3** − $7NTf_2^-$]$^{7+}$, 731.2 [**3** − $8NTf_2^-$]$^{8+}$.

### Host–guest chemistry

Cage **3** (1.0 mg, 0.12 μmol), polyaromatic guest (5 or 40 equiv.) and $CD_3CN$ (0.5 ml) were added to a 2-ml glass vial. The mixture was sonicated for 30 s and stirred at room temperature for 1 h. The formation of a host–guest complex was confirmed by NMR and ESI-MS analyses.

## Data availability

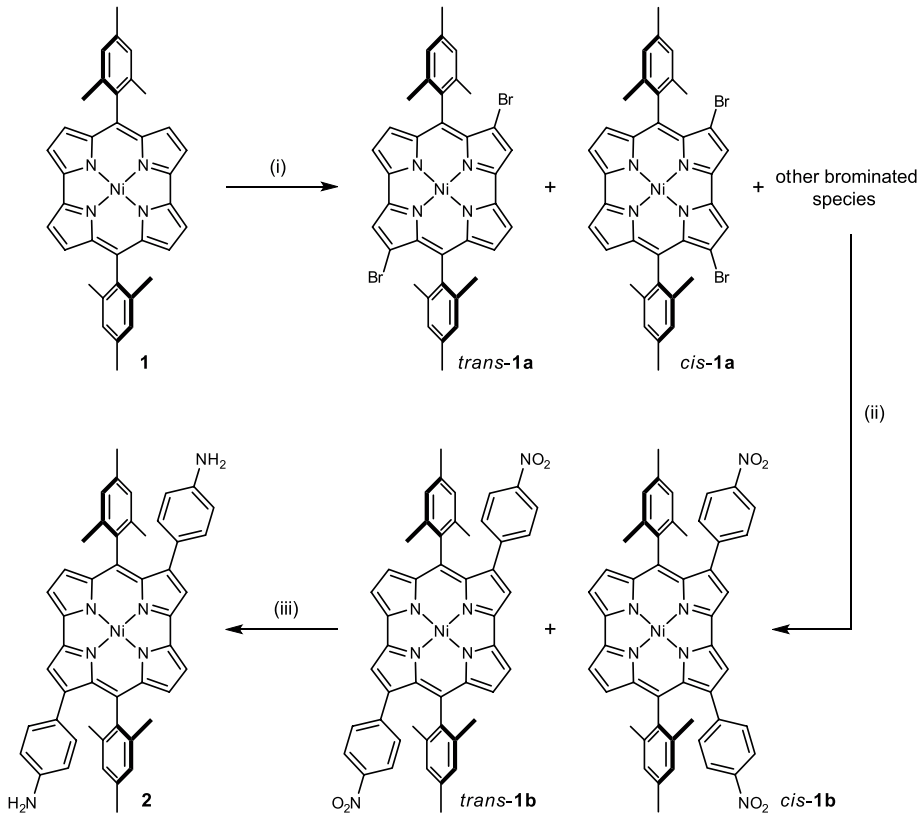
All data needed to evaluate the conclusions given in the paper are present in the paper and Supplementary Information. Any additional data related to this paper may be requested from the authors. Crystallographic data for the structure reported in this paper has been deposited at the Cambridge Crystallographic Data Centre (deposition number 1893553) and can be obtained free of charge via www.ccdc.cam.ac.uk/data_request/cif.

**Acknowledgements** This study was supported by the European Research Council (695009) and the UK Engineering and Physical Sciences Research Council (EPSRC, EP/P027067/1). M.Y. acknowledges the Japan Society for the Promotion of Science (JSPS) for an Overseas Research Fellowship. R.L. was funded by a Fondation Wiener-Anspach postdoctoral fellowship. M.P. acknowledges support from the Danish Council for Independent Research (DFF 4181-00206). We thank the NMR service of the University of Cambridge Chemistry Department for NMR experiments. We thank Diamond Light Source for providing time on Beamline I19 (MT15768), and D. Allan and L. Saunders for assistance. The computations were performed at the Research Center for Computational Science, Okazaki, Japan. We appreciate discussions with H. Shinokubo (Nagoya University), M. Yoshizawa (Tokyo Institute of Technology) and T. Soya (Kyoto University).

**Author contributions** M.Y. and J.R.N. designed the work, carried out research, analysed data and wrote the paper. Y.T. contributed to theoretical calculations. R.L. contributed to data analyses and calculations. T.K.R. contributed to X-ray crystallographic analysis. M.P. synthesized tetraoxa[8]circulene. J.R.N. is the principal investigator. All authors discussed the results and commented on the manuscript.

**Competing interests** The authors declare no competing interests.

**Additional information**
**Supplementary information** is available for this paper at https://doi.org/10.1038/s41586-019-1661-x.
**Correspondence and requests for materials** should be addressed to J.R.N.




**Extended Data Fig. 1 | Preparation of diamine 2 from norcorrole 1.** When **1** was treated with 1,3-dibromo-5,5-dimethylhydantoin, inseparable mono-, di- and tribrominated products were obtained. As bromination randomly occurred at positions 3, 7, 12 and 16 (that is, adjacent to the mesityl moieties), the 3,12- and 3,16-substituted dibromonorcorroles *trans*-**1a** and *cis*-**1a** were obtained as the main products (Supplementary Figs. 1, 2). Subsequently, disubstituted *trans*-**1b** and *cis*-**1b** were obtained by Suzuki–Miyaura cross-coupling at 43% yield as a mixture of two regioisomers (Supplementary Figs. 3, 4). Finally, 3,12-substituted subcomponent **2** could be isolated as a single regioisomer at 46% yield following reduction of the $NO_2$ groups and chromatography (Supplementary Figs. 5–9). Reagents and conditions: (i) 1,3-dibromo-5,5-dimethylhydantoin, $CH_2Cl_2$, –78 °C, 1 h, 94% (mixture of *trans*-**1a**, *cis*-**1a** and other brominated species). (ii) 4-Nitrophenylboronic acid pinacol ester, $Pd(PPh_3)_4$, $K_3PO_4$, dry THF, 70 °C, 1 h, 43% (mixture of *trans*-**1b** and *cis*-**1b**). (iii) $SnCl_2 \cdot H_2O$, dry THF, 70 °C, overnight, 46% (isolated **2**).

# Site-specific allylic C–H bond functionalization with a copper-bound N-centred radical

https://doi.org/10.1038/s41586-019-1655-8

Received: 27 April 2019

Accepted: 14 August 2019

Published online: 23 October 2019

Jiayuan Li[1], Zhihan Zhang[2], Lianqian Wu[1], Wen Zhang[1], Pinhong Chen[1], Zhenyang Lin[2]* & Guosheng Liu[1]*

Methods for selective C–H bond functionalization have provided chemists with versatile and powerful toolboxes for synthesis, such as the late-stage modification of a lead compound without the need for lengthy de novo synthesis[1–5]. Cleavage of an $sp^3$ C–H bond via hydrogen atom transfer (HAT) is particularly useful, given the large number of available HAT acceptors and the diversity of reaction pathways available to the resulting radical intermediate[6–17]. Site-selectivity, however, remains a formidable challenge, especially among $sp^3$ C–H bonds with comparable properties. If the intermediate radical could be further trapped enantioselectively, this should enable highly site- and enantioselective functionalization of C–H bonds. Here we report a copper (Cu)-catalysed site- and enantioselective allylic C–H cyanation of complex alkenes, in which a Cu(II)-bound nitrogen (N)-centred radical plays the key role in achieving precise site-specific HAT. This method is shown to be effective for a diverse collection of alkene-containing molecules, including sterically demanding structures and complex natural products and pharmaceuticals.

To achieve site-selective hydrogen atom abstraction, strategies typically rely on substrate control, whereby C–H bonds with different bond dissociation energies, polarities, steric or electronic factors dictate the site-selectivity[6–8,12–14]. Organic molecules often contain multiple $sp^3$ C–H bonds having comparable properties, and therefore exhibit undistinguishable reactivity with reagents capable of promoting HAT pathways (small $\Delta\Delta G^{\ddagger}$, Fig. 1a)[10,11]. Our goal was to increase the intrinsic reactivity differences among different C–H bonds, thereby enabling HAT with enhanced site-selectivity. To this end, we show that a Cu(II)-bound N-centred radical (NCR), with a modular sulfonamide moiety and bidentate ligand (Fig. 1b), acts as a tunable HAT reagent capable of amplifying the site-selectivity among similar allylic C–H bonds in complex molecules (Fig. 1c). Subsequent regio-, stereo- and enantioselective capture of the allylic radical by chiral Cu(II)-cyanide species leads to highly selective allylic C–H cyanation (Fig. 1c). This strategy provides a powerful tool for late-stage functionalization of complex alkene-containing molecules, including natural products and drugs.

Alkenes represent an important class of functional groups in fine chemicals, natural products, pharmaceuticals and organic materials. A long-sought-after organic transformation by synthetic chemists is thus the asymmetric functionalization of allylic C–H bonds, which would grant access to highly valuable, optically pure olefinic compounds from readily available alkenes[18,19]. Of particular importance is late-stage functionalization of olefin-containing bioactive molecules, which would provide a straightforward way to efficiently construct drug candidate libraries. However, after nearly 60 years of development, asymmetric Kharasch–Sosnovsky reactions[20–22] still suffer from limited substrate scope (only simple cyclic alkenes) and the requirement of excess alkenes, not to mention challenges of site-selectivity (Fig. 1c).

We recently developed a Cu-catalysed radical relay as an effective approach for the asymmetric cyanation of styrenes[23] and benzylic C–H bonds[24]. The benzylic radical selectively captured by a chiral Cu(II) cyanide is the key step to constructing a C–CN bond with excellent enantiomeric excess. We were delighted to find that the allylic radical can also be regioselectively and enantioselectively trapped by chiral Cu(II) cyanide to deliver allylic cyanation products (see Supplementary Fig. 1).

To survey the possible site-selective HAT, the allylic cyanation of trisubstituted alkenes **1a** bearing two sets of allylic hydrogens was performed (Fig. 2a). A moderate site-selectivity (C3:C7 = 5:1) was obtained with *N*-fluoroalkylsulfonamide (NFAS) **2a** to give **3a** at 13% yield, where NFAS acts as the precursor of structurally diverse NCRs. Varying substituents on NFAS greatly increased the selectivity and efficiency of the reaction. For instance, both site-selectivity and yield were greatly improved by introducing a bulky alkyl group on the nitrogen atom (from **2a** at 13% yield, C3:C7 = 5:1 to **2d** at 67% yield, C3:C7 = 17:1). The site-selectivity was further improved by introducing an electron-withdrawing aryl group (**2e**, C3:C7 = 22:1). Pleasingly, all of the reactions (Fig. 2a) produced the stereoisomer **3a** with excellent enantiomeric excess values, indicating that the allylic radical generated was captured by a chiral Cu(II) cyanide in a highly regio- and enantioselective manner. Similar enantiomeric excess values (89–91%) were obtained with different NCR precursors **2a** to **2e**, confirming that the enantioselective radical trapping process by the Cu species was independent of the hydrogen atom abstraction step (Fig. 1c).

[1]State Key Laboratory of Organometallic Chemistry, Center for Excellence in Molecular Synthesis, Shanghai Institute of Organic Chemistry, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai, China. [2]Department of Chemistry, The Hong Kong University of Science and Technology, Hong Kong, China. *e-mail: chzlin@ust.hk; gliu@mail.sioc.ac.cn

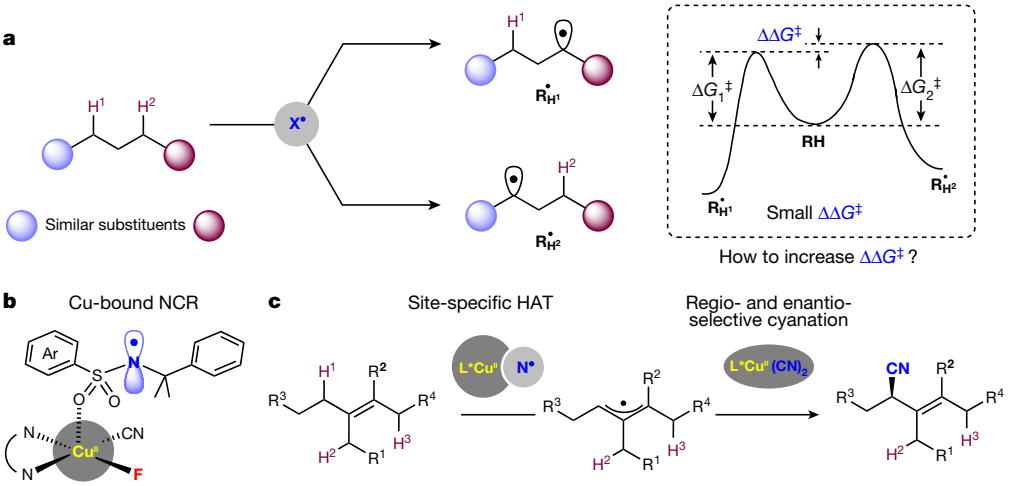



**Fig. 1 | Site- and enantioselective oxidation of $sp^3$ C–H bonds. a**, HAT of comparable C–H bonds and the corresponding energy diagram. **b**, Proposed Cu(II)-bound NCR. **c**, Site-specific and enantioselective allylic C–H cyanation of multi-substituted alkenes (this work). R, radical; L, ligand.

More importantly, the choice of ligand had a remarkable impact on the site-selectivity in that a lower site-selectivity was observed with ligand **L2** compared to **L1**. As shown in Fig. 2a, the site-selectivity (**3a**, C3:C7) was dramatically decreased from 22:1 to 5:1. The influence by both NFAS and the ligand on the site-selectivity (C3:C7) suggests that the HAT process is not simply controlled by a free NCR. An electrophilic *N*-xanthylamide reagent has recently been used to enable $sp^3$ C–H xanthylation[25], where free amidyl radicals were involved in the removal of hydrogen atoms[26,27]. Inspired by this chemistry, we applied the similar reagent **4**, which contains the same sulfonamide moiety as the electrophilic NFAS (N–F reagent) **2e**, to probe the site-selectivity of the reaction of **1a** under a free radical pathway (Fig. 2b). The reactions gave a much lower site-selectivity (**5a**, C3:C7 = 3.4:1) than a Cu-catalysed system (**3a**, C3:C7 = 22:1), highlighting the limitations of using free NCRs as HAT acceptors for site-selective C–H functionalization. High site-selectivity was also observed in the asymmetric allylic cyanation of non-conjugated acyclic alkene **1b** (Fig. 2c), selectively abstracting allylic hydrogen at C3 to give **3b** with good reactivity (64% yield) and site-selectivity (C3:C6:C7 = 10:1:0.6). By contrast, the radical chain pathway gave xanthylation product **5b** with a poor selectivity (C3:C6:C7 = 2:1:0.5). These results show that the NCRs involved in the allylic cyanation are likely to bind to Cu.

The proposed Cu(II)-bound NCRs are also supported by a number of experiments. First, the NFAS reagent **2f** was employed to test for possible asymmetric radical cyclization. The enantioselective induction was observed with the chiral ligand **L3**, giving product **6** in 92% yield with 15% enantiomeric excess (Fig. 3a). Second, the oxidation of (**L1**)Cu(I) with the N–F reagent **2e** was monitored by electron paramagnetic resonance spectroscopy, which showed a Cu(II) signal (Fig. 3b (i)). An NCR was also probed by adding a radical trap 5,5-dimethyl-1-pyrroline-*N*-oxide (DMPO) **7** (Fig. 3b (ii)).

Density functional theory calculations at the M06 level of theory were performed to gain more mechanistic insights. The reaction of (**L1**)Cu(I)(CN) with the N–F reagent **2e** was first calculated, showing that the reaction readily gave a Cu(II)-bound NCR species (Fig. 1b and Fig. 3c). This Cu(II)-bound NCR species is more stable by 9.4 kcal mol$^{-1}$ in free energy than (**L1**)Cu$^{II}$(CN)F + free NCR. For comparison, other possible species were also calculated, but are less stable than the Cu(II)-bound NCR via a Cu–O coordination (see Supplementary Fig. 9). The activation barriers for HAT of allylic C–H bonds of **1a** by Cu(II)-bound NCR were then calculated (path b), and compared with those by the free NCR (path a). The results presented in Fig. 3c reveal that, upon coordination of the NCR to Cu(II), the barriers associated with HAT increase from 8.1 kcal mol$^{-1}$ to 12.8 kcal mol$^{-1}$ and from 9.6 kcal mol$^{-1}$ to 15.4 kcal mol$^{-1}$, respectively. The corresponding $\Delta\Delta G^{\ddagger}$ for allylic HAT from sites C3 and C7 increases from 1.5 kcal mol$^{-1}$ to 2.6 kcal mol$^{-1}$ (that is, $\Delta\Delta\Delta G^{\ddagger}$ = 1.1 kcal mol$^{-1}$)[28], leading to an increase of nearly an order of magnitude in the site-selectivity. This computational outcome closely resembles the experimental results obtained from the reaction of **1a** (compare C3:C7 = 3.4:1 in Fig. 2b to C3:C7 = 22:1 with NFAS reagent **2e** in Fig. 2a, $\Delta\Delta\Delta G^{\ddagger}$ = 1.2 kcal mol$^{-1}$). In short, O-coordination of NCR to the Cu(II) centre leads to a reactive radical that exhibits greater selectivity in its reaction with C–H bonds relative to the parent NCR. To our knowledge, this catalytic modulation of radical reactivity is unique and complements previous efforts to modulate reactivity by using different stoichiometric reagents[29].

To explore the scope of the site- and enantioselective allylic C–H cyanation reaction, two NCR precursors **2d** and **2e** (Fig. 2a) were applied to a wide range of alkenes (Fig. 4a (i)). Internal alkenes—including vinyl(hetero)arenes (**1c** to **1e**), enyne (**1f**), vinylboronic ester (**1g**) and allylic imide (**1h**)—proved to be viable substrates, providing the desired *E*-stereoisomers **3c** to **3h** in good yields (42–91%) with excellent regio- (>20:1) and enantioselectivity (87–96% enantiomeric excess). Various functional groups (such as phthalimide, silyl and boronic ester) and heterocycles (such as furan and pyridine) can be tolerated (for more examples, see Supplementary Fig. 2). Notably, some enantiomerically enriched allylic nitriles (*E*-**3d** and *E*-**3e**) were efficiently obtained from the mixtures of *E*- and *Z*-isomeric alkenes synthesized from the Wittig reaction. The practical utility of our method in the laboratory was demonstrated by the gram-scale preparations of **3c** and **10b** (Fig. 4b).

Our attention then turned to trisubstituted alkenes bearing multiple sets of allylic hydrogens, which are challenging in terms of both reactivity and site-specificity (Fig. 4a (ii)). Excellent site-selectivity was observed for acyclic trisubstituted alkenes **1i** to **1m**, in which similar allylic hydrogens at the C3 and C4 positions are differentiated by the Cu(II)-bound NCR generated from **2e** and (**L1**)Cu(I). Various functional groups—such as vinylsilane (*E*-**1k**), vinylimides (*E*-**1l**) and vinylester (*Z*-**1m**)—are compatible with the standard reaction conditions, providing the corresponding allylic nitriles **3k** to **3m** with both excellent site-selectivity (>20:1) and enantioselectivity (89–95% enantiomeric excess). Notably, both isomers *E*-**1k** and *Z*-**1k** exhibited similar reactivity to give the same product *E*-**3k** in 67–72% yields with 95% enantiomeric excess, which can be applied to the synthesis of enantiomerically enriched internal allylic nitriles (**11**) via desilylation (Fig. 4c (i)). The selective allylic C–H cyanations were also successful with tri-alkyl substituted alkenes having three sets of allylic hydrogens (C1, C3 and C4), such as allylic ester and allylic imide, to give single isomers *E*-**3n** (60% yield, 81% enantiomeric excess) and *E*-**3o** (52% yield, 85% enantiomeric excess). Moreover, the Cu(II)-NCR also shows exceptional sensitivity for **1p**, with one methylene (C3) and two methyl groups (C6, C7) in the allylic positions (bond dissociation

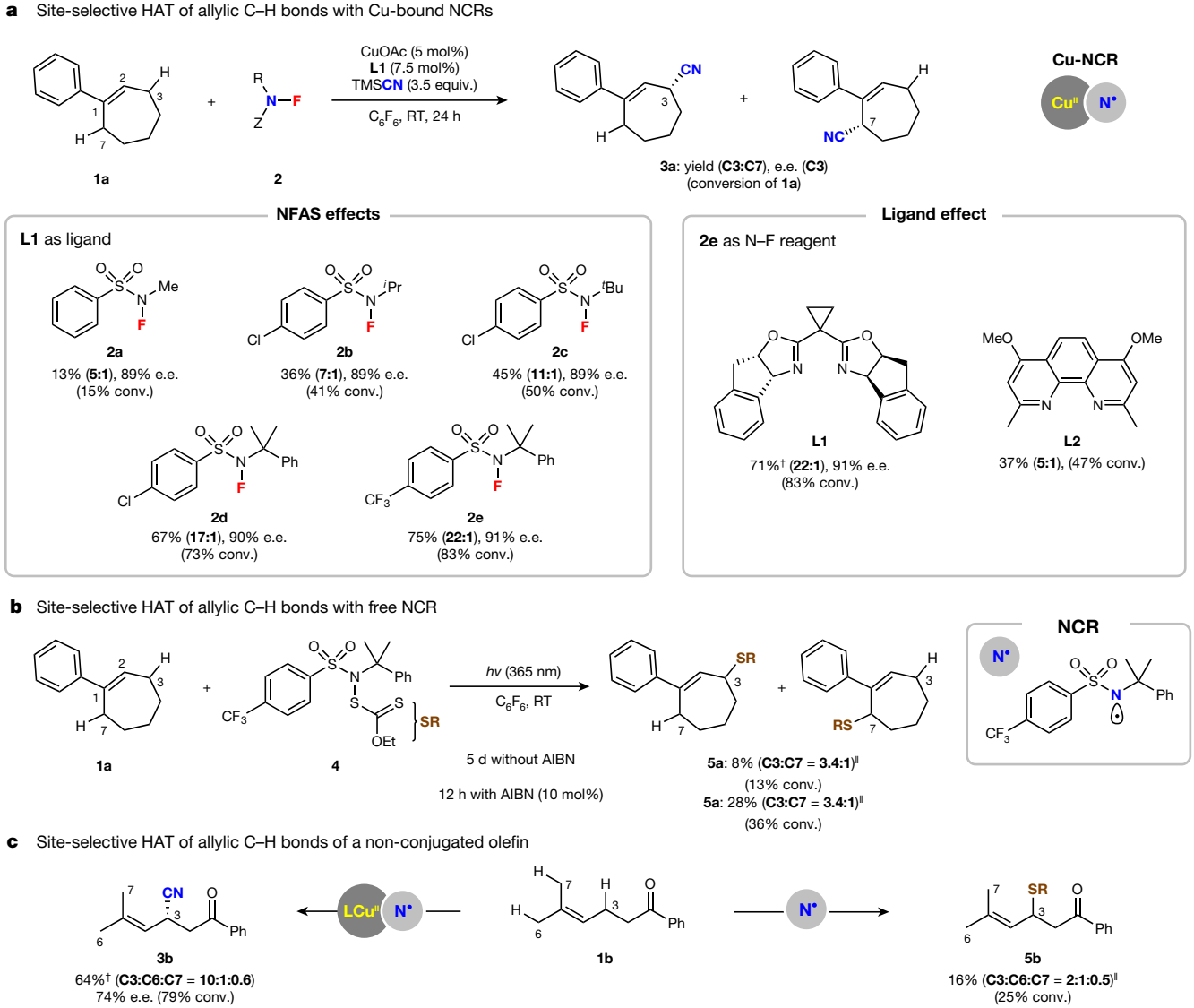



**Fig. 2 | Enantioselective allylic cyanation of alkenes. a**, Optimization of reaction conditions: **1a** (0.2 mmol), N–F reagent **2** (0.6 mmol), trimethylsilyl cyanide (TMSCN, 0.7 mmol), CuOAc (5 mol%), ligands **L1** and **L2** (7.5 mol%) in $C_6F_6$ (1.0 ml) at room temperature (RT; 20–25 °C) for 24 hours. Yield and site-selectivity ratio were determined by $^1$H nuclear magnetic resonance (NMR) spectroscopy. The enantiomeric excess (e.e.) value was determined by high-performance liquid chromatography (HPLC). †The isolated yield of major product **3a** or **3b** (C3; the product with the cyano group in the carbon-3 position). **b**, Site-selective HAT by free NCR. *hv*, ultraviolet light. ‖The isolated yield of the mixture of regiomers **5a** or **5b**. **c**, Site-selective HAT of alkene **1b** with two distinct HAT acceptors. AIBN, azobisisobutyronitrile; 'conv.' refers to conversion of the substrate.

energy ~83 kcal mol$^{-1}$) and one benzylic hydrogen (C3′, bond dissociation energy ~85 kcal mol$^{-1}$), and abstracts the allylic hydrogen at C3 selectively (C3:C6:C7:C3′ = 10:1:0.6:0). Furthermore, the unconjugated diene **1q** (*E*:*Z* = 1:1) with two similar steric allylic methylene moieties showed good site-selectivity for the C3 position (C3:C5 = 9:1). As examples of substrates with a cyclic structural motif, **1r** and **1s** underwent excellent site-specific and enantioselective cyanation (91–99% enantiomeric excess). The bulkier tetrasubstituted alkene **1t**, usually inert towards metal catalysis, was reactive with excellent site-selectivity to afford product **3t** in moderate yield and good enantioselectivity. These results showcase that the exquisite site-selectivity of HAT by a Cu(II)-bound NCR is governed by very subtle differences in steric environments.

For the optically pure allylic ester (*R*)-*E*-**1u**, catalysts with either enantiomer of ligand **L1** were employed to probe stereoselectivity. Both led to excellent site-selectivity (>20:1) and good to excellent diastereoselectivity (81–91%), indicating high levels of catalyst-controlled rather than substrate-controlled stereoselectivity (Fig. 4a (iii)). The excellent functional group compatibility and site-specificity enabled us to synthesize enantiomerically enriched compound **8**, an inhibitor of dipeptidyl peptidase IV[30], via highly selective allylic C–H cyanation of alkene (*S*,*S*)-**1v** followed by deprotection of the NHBoc group (Fig. 4a (iii)).

The radical relay method reported here is also applicable for highly selective late-stage functionalization of bioactive compounds, such as the commercial drugs altrenogest acetate, norethisterone acetate, genipin, brefeldin A and cyclefenil (Fig. 4b). The reaction of altrenogest acetate **9a**, containing four different types of allylic hydrogen, exhibited remarkably high site-, regio- and diastereoselectivity to provide only one isomer **10a** in 65% yield, along with the recovered **9a** (22%). Norethisterone acetate **9b**, which contains three different types of allylic hydrogens along with a terminal alkyne moiety, afforded the single isomer **10b** selectively, with the terminal alkyne moiety intact. Diacetates of genipin **9c** and brefeldin A **9d**, which bear three or four sets of allylic hydrogens along with abstractable hydrogens adjacent to ester groups, exhibited excellent site-specificity and diastereoselectivity,

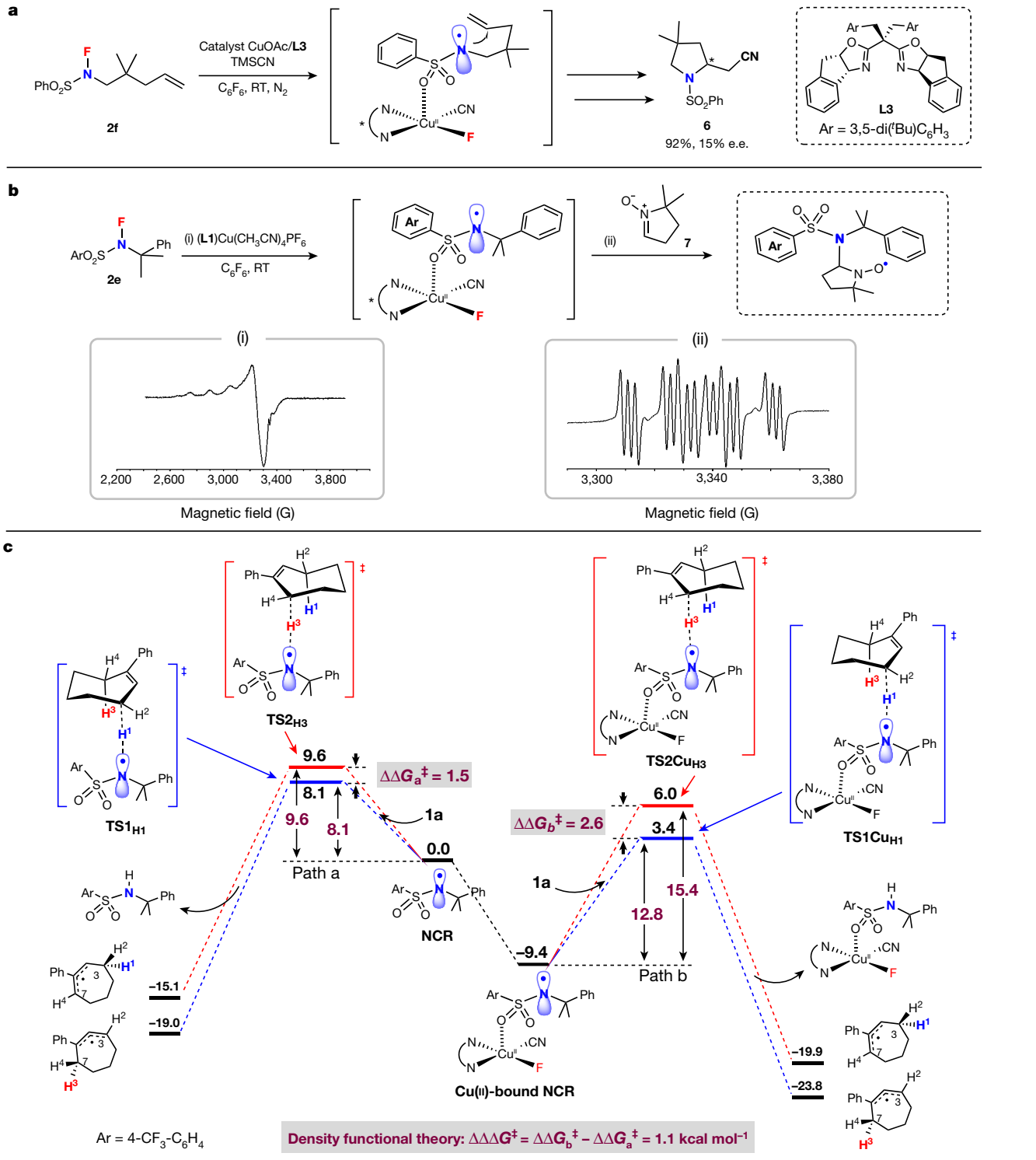



**Fig. 3 | Mechanistic studies. a**, Asymmetric radical cyclization of chiral Cu(II)-bound NCRs. **b**, Electron paramagnetic resonance studies. (i) Cu(II) signal of Cu(II)-bound NCR; (ii) Cu(II)-bound NCR adding to DMPO (**7**). **c**, Results of density functional theory calculations (all values are in kilocalories per mole) at the Minnesota 2006 functional (M06) level of theory. TS1 and TS2 are transition states. We note that (1) the Cu(II)-bound NCR species is more stable by 9.4 kcal mol$^{-1}$ in free energy than (**L1**)Cu(II)(CN)F + free NCR, and (2) C3/C7 each has two CH bonds and only the favourable HAT processes of $H^1$ and $H^3$ are presented here (see Supplementary Figs. 11, 12).

leading to single cyanation products **10c** and **10d**, respectively. Phytol, a natural product with a linear chain, was converted to ether **9e**, which contains three sets of allylic C–H hydrogens and three tertiary C–H bonds. The reaction gave allylic cyanation product **10e** in 51% yield and 74% diastereomeric excess, along with 5% cyanation of the methyl group (C3). Unlike the Rh-carbene-promoted C–H insertion[31] and Fe-catalysed aliphatic C–H activation[32], no tertiary C–H functionalization was observed here. Finally, as an example for a rather inert

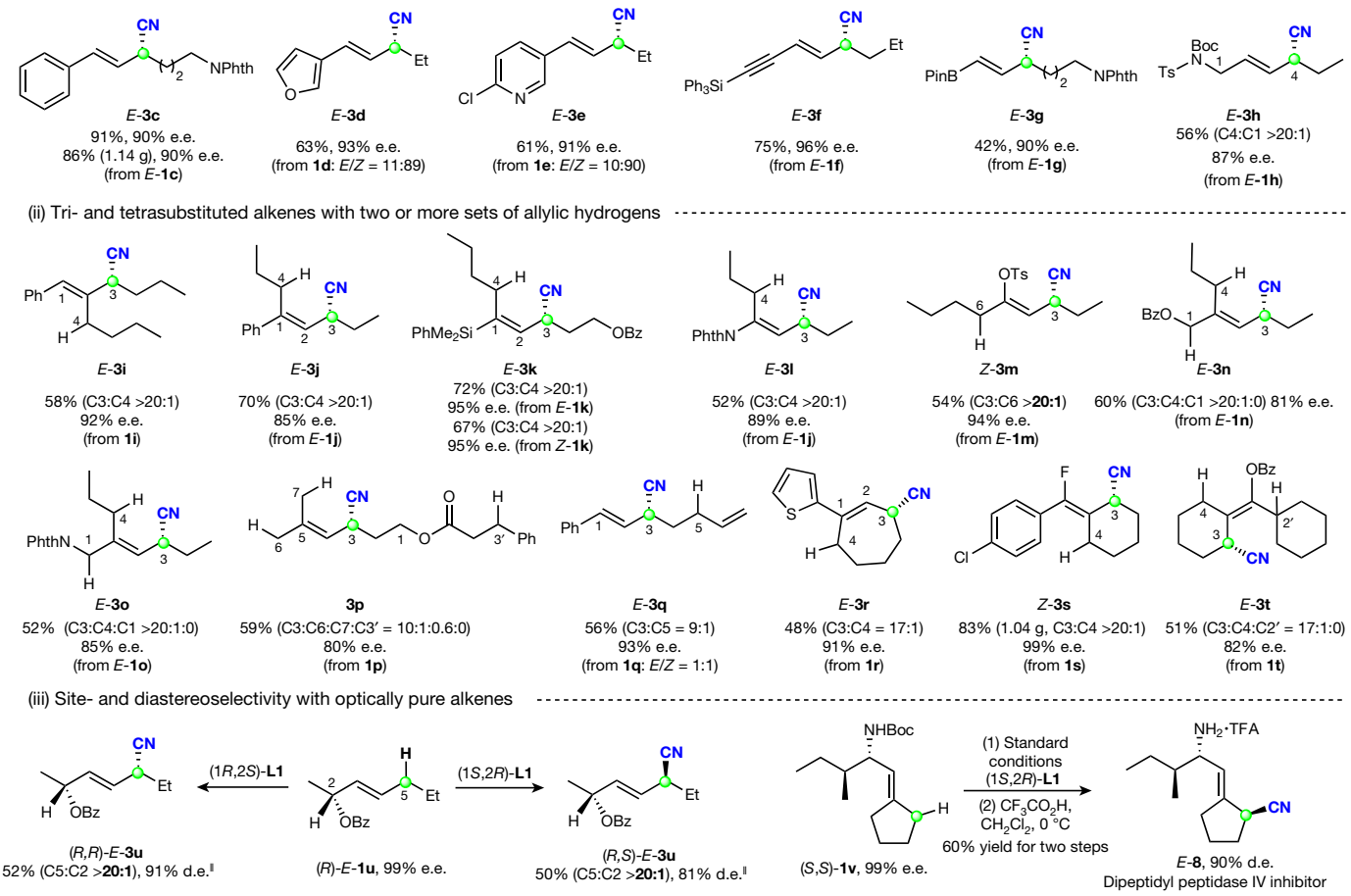



**Fig. 4 | Substrate scope and late-stage cyanation of drug derivatives. a**, Substrate scope. **b**, Late-stage cyanation of drugs. #All prepared with (1*R*,2*S*)-**L1**. **c**, Synthetic applications. See Supplementary Information for experimental details; the isolated yield of major regiostereomer or diastereomer, and the site-selectivity was determined by $^1$H NMR of the crude mixture. †The regioselectivity (**3c** to **3h**) >20:1. ‖The diastereomic excess (d.e.) value of the product was determined by HPLC. $Single isomer (diastereoselectivity d.r. > 20:1) was observed in crude $^1$H NMR. ¶*E*-**11**:*Z*-**11** = 10:1, §**12a** (d.r. = 9:1), **15a** (d.r. = 9:1), **15b** (d.r. = 6:1). (r.s.m. is recovered starting material.) See the Supplementary Information for details of the Shi-epoxidation and the Raney-Ni reduction; 'standard conditions' refers to those used in Fig. 2a with NFAS **2e**.

tetrasubstituted olefin, a drug named cyclefenil **9f** was reactive to give **10f** in 43% yield with 93% enantiomeric excess.

Optically active nitriles are of great importance to the pharmaceutical industry owing to their excellent biocompatibilities, resistance to metabolism, and abilities to accentuate biological activities through hydrogen-bonding interactions[33]. Results in Fig. 4a show that the radical relay method reported here represents the most efficient method to date for the synthesis of various optically active allylic nitriles from readily available olefins. Additionally, these enantiomerically enriched nitriles can serve as synthetically valuable building blocks to optically active alkyl amines and carboxamides (Fig. 4c (ii) to (iv)). For instance, allylic nitrile *Z*-**3s** can be efficiently converted to alkyl nitrile **12a** and homoallylic amine derivative **12b** via selective hydrogenation with the stereocentres preserved. It should be noted that the chiral alkyl nitrile **13** was obtained from allylic C–H cyanation of the *E*/*Z* mixture of alkenes **1w** and subsequent hydrogenation in one pot. Hydrolysis converts **13** to the corresponding chiral alkyl carboxamide **14** efficiently. In addition, the allylic nitrile **3c** was subjected to epoxidation to give chiral epoxide **15a** in good diastereoselectivity (9:1), which in turn underwent regioselective ring-opening hydrofluorination to give the highly functionalized chiral nitrile **15b** efficiently.

This reaction illustrates that the Cu-based radical relay is a powerful strategy for direct enantioselective functionalization of complex alkenes. Our method of site-specific allylic HAT by Cu(II)-bound NCRs lays the groundwork for further exploration of site-specific and enantioselective C–H oxidation reactions, including those with allylic and other types of C–H bonds.

## Data availability

The data supporting the findings of this study are available within the paper and its Supplementary Information. Metrical parameters for the structure of reagents **2c**, **2d** and products **10d**, **17c** and **S8** (see Supplementary Information) are available free of charge from the Cambridge Crystallographic Data Centre (https://www.ccdc.cam.ac.uk/) under reference numbers CCDC 1945032, CCDC 1945034, CCDC 1945035, CCDC 1945033 and CCDC 1945036, respectively.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1655-8.


1. Cernak, T., Dykstra, K. D., Tyagarajan, S., Vachal, P. & Krska, S. W. The medicinal chemist's toolbox for late stage functionalization of drug-like molecules. *Chem. Soc. Rev.* **45**, 546–576 (2016).
2. Davies, H. M. L. & Morton, D. Guiding principles for site-selective and stereoselective intermolecular C–H functionalization by donor/acceptor rhodium carbenes. *Chem. Soc. Rev.* **40**, 1857–1869 (2011).
3. Zhang, F.-L., Hong, K., Li, T.-J., Park, H. & Yu, J.-Q. Functionalization of C($sp^3$)-H bonds using a transient directing group. *Science* **351**, 252–256 (2016).
4. McNally, A., Haffemayer, B., Collins, B. S. L. & Gaunt, M. J. Palladium-catalysed C–H activation of aliphatic amines to give strained nitrogen heterocycles. *Nature* **510**, 129–133 (2014).
5. Prier, C. K., Zhang, R. K., Buller, A. R., Brinkmann-Chen, S. & Arnold, F. H. Enantioselective, intermolecular benzylic C–H amination catalyzed by an engineered iron-haem enzyme. *Nat. Chem.* **9**, 629–634 (2017).
6. Sharma, A. & Hartwig, J. F. Metal-catalyzed azidation of tertiary C–H bonds suitable for late-stage functionalization. *Nature* **517**, 600–604 (2015).
7. Le, C., Liang, Y., Evans, R. W., Li, X. & MacMillan, D. W. C. Selective $sp^3$ C–H alkylation via polarity-match-based cross-coupling. *Nature* **547**, 79–83 (2017).
8. Chen, M. S. & White, M. C. Combined effects on selectivity in Fe-catalyzed methylene oxidation. *Science* **327**, 566–571 (2010).
9. Horn, E. J. et al. Scalable and sustainable electrochemical allylic C–H oxidation. *Nature* **533**, 77–81 (2016).
10. Liu, W. & Groves, J. T. Manganese catalyzed C–H halogenation. *Acc. Chem. Res.* **48**, 1727–1735 (2015).
11. Che, C.-M., Lo, V. K.-Y., Zhou, C.-Y. & Huang, J.-S. Selective functionalisation of saturated C–H bonds with metalloporphyrin catalysts. *Chem. Soc. Rev.* **40**, 1950–1975 (2011).
12. Jeffrey, J. L., Terrett, J. A. & MacMillan, D. W. C. O–H hydrogen bonding promotes H-atom transfer from a C–H bonds for α-alkylation of alcohols. *Science* **349**, 1532–1536 (2015).
13. Schmidt, V. A., Quinn, R. K., Brusoe, A. T. & Alexanian, E. J. Site-selective aliphatic C–H bromination using *N*-bromoamides and visible light. *J. Am. Chem. Soc.* **136**, 14389–14392 (2014).
14. Milan, M., Salamone, M., Costas, M. & Bietti, M. The quest for selectivity in hydrogen atom transfer based aliphatic C–H bond oxygenation. *Acc. Chem. Res.* **51**, 1984–1995 (2018).
15. Carestia, A. M., Ravelli, D. & Alexanian, E. J. Reagent-dictated site selectivity in intermolecular aliphatic C–H functionalizations using nitrogen-centered radicals. *Chem. Sci.* **9**, 5360–5365 (2018).
16. Shields, B. J. & Doyle, A. G. Direct C($sp^3$)–H cross coupling enabled by catalytic generation of chlorine radicals. *J. Am. Chem. Soc.* **138**, 12719–12722 (2016).
17. Milan, M., Bietti, M. & Costas, M. Highly enantioselective oxidation of nonactivated aliphatic C–H bonds with hydrogen peroxide catalyzed by manganese complexes. *ACS Cent. Sci.* **3**, 196–204 (2017).
18. Covell, D. J. & White, M. C. A chiral Lewis acid strategy for the enantioselective allylic C–H oxidation. *Angew. Chem. Int. Ed.* **47**, 6448–6451 (2008).
19. Bayeh, L., Le, P. Q. & Tambar, U. K. Catalytic allylic oxidation of internal alkenes to a multifunctional chiral building block. *Nature* **547**, 196–200 (2017).
20. Kharasch, M. S. & Sosnovsky, G. The reactions of *t*-butyl perbenzoate and olefins—a stereospecific reaction. *J. Am. Chem. Soc.* **80**, 756 (1958).
21. Eames, J. & Watkinson, M. Catalytic allylic oxidation of alkenes using an asymmetric Kharasch–Sosnovsky reaction. *Angew. Chem. Int. Ed.* **40**, 3567–3571 (2001).
22. Andrus, M. B. & Zhou, Z. Highly enantioselective copper-bisoxazoline-catalyzed allylic oxidation of cyclic olefins with tert-butyl *p*-nitroperbenzoate. *J. Am. Chem. Soc.* **124**, 8806–8807 (2002).
23. Wang, F. et al. Enantioselective copper-catalyzed intermolecular cyanotrifluoromethylation of alkenes via radical process. *J. Am. Chem. Soc.* **138**, 15547–15550 (2016).
24. Zhang, W. et al. Enantioselective cyanation of benzylic C–H bonds via copper-catalyzed radical relay. *Science* **353**, 1014–1018 (2016).
25. Czaplyski, W. L., Na, C. G. & Alexanian, E. J. C–H xanthylation: a synthetic platform for alkane functionalization. *J. Am. Chem. Soc.* **138**, 13854–13857 (2016).
26. Kärkäs, M. D. Photochemical generation of nitrogen-centered amidyl, hydrazonyl, and imidyl radicals: methodology development and catalytic applicants. *ACS Catal.* **7**, 4999–5022 (2017).
27. Artaryan, A. et al. Aliphatic C–H bond iodination by a *N*-iodoamide and isolation of an elusive *N*-amidyl radical. *J. Org. Chem.* **82**, 7093–7100 (2017).
28. Palucki, M. et al. The mechanistic basis for electronic effect on enantioselectivity in the (salen)Mn(II)-catalyzed epoxidation reaction. *J. Am. Chem. Soc.* **120**, 948–954 (1998).
29. Barham, J. P., John, M. P. & Murphy, J. A. Contra-thermodynamic hydrogen atom abstraction in the selective C–H functionalization of trialkylamine *N*-$CH_3$ groups. *J. Am. Chem. Soc.* **138**, 15482–15487 (2016).
30. Peters, J.-U. 11 years of cyanopyrrolidines as DPP-IV inhibitors. *Curr. Top. Med. Chem.* **7**, 579–595 (2007).
31. Liao, K. et al. Site-selective and stereoselective functionalization of non-activated tertiary C–H bonds. *Nature* **551**, 609–613 (2017).
32. Chen, M. S. & White, M. C. A predictably selective aliphatic C–H oxidation reaction for complex molecule synthesis. *Science* **318**, 783–787 (2007).
33. Fleming, F. F., Yao, L., Ravikumar, P. C., Funk, L. & Shook, B. C. Nitrile-containing pharmaceuticals: efficacious roles of the nitrile pharmacophore. *J. Med. Chem.* **53**, 7902–7917 (2010).

**Acknowledgements** Financial support was provided by the National Basic Research Program of China (grant number 973-2015CB856600), the National Natural Science Foundation of China (grant numbers 21532009, 21821002, 21790330 and 21761142010), the Science and Technology Commission of Shanghai Municipality (grant numbers 17XD1404500, 17QA1405200 and 17JC1401200), the Strategic Priority Research Program (grant number XDB20000000), the Key Research Program of Frontier Science (grant number QYZDJSSW-SLH055) and the Youth Innovation Promotion Association (grant number 2018292) of the Chinese Academy of Sciences, and the Research Grants Council of Hong Kong (grant numbers HKUST 16304416 and 16304017). J.L. thanks Y. Zhang for performing the electron paramagnetic resonance manipulation and analysis and X. Wan for HPLC analysis. G.L. thanks H. Guan (University of Cincinnati) and S. Stahl (University of Wisconsin-Madison) for discussions.

**Author contributions** J.L., L.W., W.Z. and G.L. conceived the work and designed the experiments. J.L. performed most of the laboratory experiments (with help from L.W. and W.Z.). Z.Z. and Z.L. conducted density functional theory calculations. J.L., P.C., Z.L. and G.L. analysed the data and wrote the manuscript.

**Competing interests** The authors declare no competing interests.

**Additional information**

**Supplementary information** is available for this paper at https://doi.org/10.1038/s41586-019-1655-8.

**Correspondence and requests for materials** should be addressed to Z.L. or G.L.

**Peer review information** *Nature* thanks John Murphy, Peter Richard Schreiner and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Article

# Similar scaling laws for earthquakes and Cascadia slow-slip events

https://doi.org/10.1038/s41586-019-1673-6

Received: 31 August 2018

Accepted: 1 August 2019

Published online: 23 October 2019

Sylvain Michel[1,2,4]*, Adriano Gualandi[1,3] & Jean-Philippe Avouac[1,5]

Faults can slip not only episodically during earthquakes but also during transient aseismic slip events[1–5], often called slow-slip events. Previous studies based on observations compiled from various tectonic settings[6–8] have suggested that the moment of slow-slip events is proportional to their duration, instead of following the duration-cubed scaling found for earthquakes[9]. This finding has spurred efforts to unravel the cause of the difference in scaling[6,10–14]. Thanks to a new catalogue of slow-slip events on the Cascadia megathrust based on the inversion of surface deformation measurements between 2007 and 2017[15], we find that a cubic moment–duration scaling law is more likely. Like regular earthquakes, slow-slip events also have a moment that is proportional to $A^{3/2}$, where $A$ is the rupture area, and obey the Gutenberg–Richter relationship between frequency and magnitude. Finally, these slow-slip events show pulse-like ruptures similar to seismic ruptures. The scaling properties of slow-slip events are thus strikingly similar to those of regular earthquakes, suggesting that they are governed by similar dynamic properties.

Geodetic monitoring of strain accumulation and release along various subduction zones has revealed episodic events of aseismic slip along different megathrusts[1–4]. These slow-slip events (SSEs) are typically accompanied by a burst of weak low-frequency seismic signals called tremors[16,17]. The characteristics of these slow earthquakes compiled from different subduction zones[6] suggest that their moment, $M_0$ (defined as the integral of slip over the fault area multiplied by the shear modulus), is proportional to their duration, $T$. It has therefore been inferred that SSEs, which obey $M_0 \propto T$, and earthquakes, which obey[9] $M_0 \propto T^3$, correspond to distinct modes of slip[6]. The cubic scaling is expected for circular ruptures with a constant stress drop expanding at a constant rate[9], a kinematic model close to the dynamic circular crack model[18] that fits most properties of earthquakes to first order. The moment–duration scaling should, however, transition to $M_0 \propto T$ for the larger 'bounded' ruptures that span the full along-dip width of the seismogenic zone[10]. This transition is hardly seen in seismicity catalogues as they are dominated by smaller, unbounded events[10,19]. By contrast, only the larger SSEs are generally detected with geodetic techniques, and they generally show strongly elongated rather than circular slipped areas, suggesting bounded ruptures. This consideration led to the suggestion[10] that the different scaling between regular earthquakes and SSEs arises because earthquake catalogues are dominated by unbounded ruptures whereas SSEs mostly represent bounded ruptures. An alternative view is that the difference of scaling between earthquakes and SSEs reflects fundamentally different dynamics[6,14].

In this study, we take advantage of a recent catalogue of SSEs from Cascadia[15], which was obtained from the inversion of geodetic position time series recorded at 352 continuous Global Positioning System (GPS) stations, from the Pacific Geodetic Array and the Plate Boundary Observatory, during the time period 2007.000 to 2017.632 (decimal year notation). After extracting a secular trend average through the SSEs from the time series, and deducing from it the pattern of coupling that represents the degree of locking of the plate interface (Fig. 1), we corrected the data for hydrological effects as well as for co-seismic and post-seismic deformation. These corrected time series were used to image spatio-temporal variations of slip along the megathrust (Fig. 1). The catalogue of SSEs extracted from the slip model history on the whole megathrust contains 64 events, most of which coincided with the spatio-temporal distribution of tremors (Fig. 2), as was found in previous similar studies[20,21]. Individual events show unidirectional or bidirectional ruptures with a rupture front velocity between about 5.5 km $d^{-1}$ and 11 km $d^{-1}$ (ref. [15]). The larger ones show pulse-like behaviour very similar to large earthquake ruptures[22] but with much lower propagation and slip rates. Figure 1 shows the cumulated distribution of slip resulting from all 64 SSEs. As shown previously[23], the zone of episodic slow slip and tremors closely follows the intersection of the forearc Moho with the megathrust and is separated from the shallower locked zone by a 40-km-wide band of steady creep (Fig. 1). The catalogue contains SSEs with a relatively wide range of sizes spanning moment magnitudes, $M_w$, between about 5.3 and 6.8 (Fig. 2), allowing the investigation of the scaling properties of a population of SSEs that all happened in a relatively narrow range of conditions.

The moment and duration of the largest events of this catalogue fall in the slow-slip domain identified by Ide et al.[6] (red shading in Fig. 3b and in Extended Data Fig. 1), but our data do not follow the linear scaling proposed in that study, and align better along the $M_0 \propto T^3$ scaling of earthquakes. This dataset, however, suffers from a bias because a low-pass temporal filter with a cut-off period of about 30 days was applied to the time series. To refine the analysis and alleviate the possibility of a bias introduced by the automatic picking of the onset and end of

[1]California Institute of Technology, Department of Geology and Planetary Sciences, Pasadena, CA, USA. [2]University of Cambridge, Department of Earth Sciences, Bullard Laboratories, Cambridge, UK. [3]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA. [4]Present address: Laboratoire de Géologie, Ecole Normale Supérieure, Paris, France. [5]Present address: Ecole Polytechnique, Palaiseau, France. *e-mail: sylvain_michel@live.fr

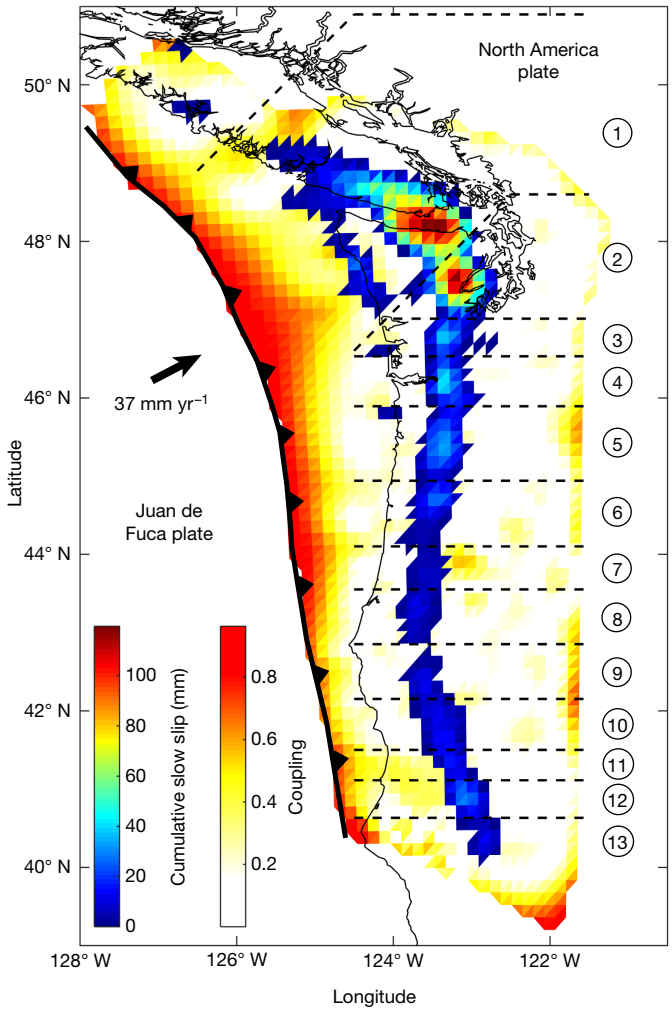



**Fig. 1 | Comparison of interseismic coupling with cumulated slip due to episodic slow slip between 2007 and 2017.** Data are from ref. [15]. The cumulated slip due to all of the 64 SSEs forms a band that follows the intersection of the forearc Moho with the megathrust and is disconnected from the shallower locked portion of the megathrust, whether the trench is assumed locked (as here) or not. Interseismic coupling, defined as the rate of slip deficit due to locking in the interseismic period divided by the long-term slip rate, and the long-term forearc motion were determined from the secular GPS velocities.

the SSEs, we carried out manual measurements using time series filtered with a shorter cut-off period of approximately 9 days (see Supplementary Information and Extended Data Fig. 2 for details). As a caution, we removed 17 events that we considered questionable, and we combined seven pairs of events owing to their closeness in time and space. The final revised catalogue consists of 40 events.

For each event, we estimate minimum and maximum durations and find the same trend as the original catalogue. We next use the revised dataset to search for the best-fitting scaling law, taking into account uncertainties in duration and magnitude, and the effect of the filter (see Methods for details). For example, we show in Fig. 3a where the filtered data should plot if $M_0 \propto T$ (orange filled circles) or $M_0 \propto T^3$ (green filled circles) were the true relationships to generate the observations. The root mean squared error (RMSE) for a scaling law exponent ($M_0 \propto T^c$) of $c = 3$ is about half the value obtained for $c = 1$ and varies little for $c \geq 3$. We also tested the robustness of our conclusion by using estimates of the SSE duration from the tremor bursts and the associated moment release based on our slip model (see Supplementary Information 'Comparison of tremor durations and SSE durations from geodesy'). In that case, no filter correction is needed. The results confirm that $M_0 \propto T^3$ is much more likely than a linear scaling, with a reduction of about 68% in the RMSE. We conclude that SSEs occurring under a narrow range of conditions (for example temperature and pressure), as is the case in the deep SSEs from Cascadia analysed here, follow a near cubic moment–duration scaling like the scaling of regular earthquakes. This finding is all the more unexpected since most of the SSEs in our catalogue ruptured the entire width of the zone of episodic slow slip (Fig. 1 and Extended Data Fig. 3) and therefore have large aspect ratios (Fig. 4). They would therefore be expected to follow a linear scaling[10]. It is noteworthy that, although the cubic scaling of regular earthquakes is generally thought to reflect self-similarity and justified on the basis of the circular crack model[24], the same scaling is observed in our dataset in which most ruptures are very elongated, with aspect ratios of 2 to 12 (Fig. 4b).

The original catalogue, as well as our manual measurement, also defines a tightly constrained scaling of moment versus rupture area, following approximately the $M_0 \propto A^{3/2}$ scaling of regular earthquakes (Fig. 3f) (the best-fitting scaling law exponent is actually 1.25; see Supplementary Information for details). The ratio $M_0 \propto A^{3/2}$ is, however, three orders of magnitude smaller, implying a stress drop of about $4.3 \pm 2.0$ kPa, based on the same circular crack model generally used to quantify seismic ruptures[9] ($M_0 = C^{-1}\Delta\tau A^{3/2}$, where $\Delta\tau$ is the stress drop, $A$ the rupture area and $C = 2.44$), compared with 1–10 MPa for regular earthquakes. This estimate of the average stress drop is, however, questionable as the rupture areas are quite elongated (Fig. 4b). We therefore estimated the average stress drop for each of our SSE based on our slip model using the approach of Noda et al.[25] and using Meade's analytical solution[26] for triangular sub-faults. The values range between 0.9 kPa and 18.0 kPa, with a mean of approximately 5.8 kPa and a standard deviation of 2.0 kPa. Our average stress drop is about 10 times lower than the value proposed by Schmidt and Gao[27] based on the slip model of 16 events with $M_w$ in the range of 6.2–6.7 between 1998 and 2008. Given that the slip distributions are similar (the three common events are compared in Extended Data Fig. 4 and in the Supplementary Information), we suspect that this difference is due to the way that rupture area was measured by Schmidt and Gao[27], the fact that our slip models do not account for the slip that would be needed to balance interseismic loading during a SSE, and the possibility that our slip models are spatially smoother because of stronger regularization of our inversion.

We also examined the frequency–magnitude scaling of the SSEs (Fig. 4a). We show the distributions obtained from both the original catalogue and the revised catalogue. The data selection in the revised catalogue results in a roll-over at lower magnitudes, but in any case we find that the SSEs approximately show a linear relationship between moment magnitude, $M_w$, and the decimal logarithm of the number of events with magnitude equal or larger than $M_w$ (that is, they obey the Gutenberg–Richter law) with a slope (called the $b$ value) of about 0.8. The abrupt drop in the frequency of events larger than $M_w = 6.4$ suggests a truncation effect. The truncation cannot be explained by the transition from unbounded ruptures to bounded ruptures in width as this transition would occur at a much lower magnitude, $M_w \approx 5.7$, given the aspect ratio of the ruptures. It could be due, instead, to the along-strike segmentation discussed below. As there are only 11 events with $M_w > 6.4$, however, this observation should be considered with caution. A previous study had also argued for SSEs obeying the Gutenberg–Richter law[28] but used moment inferred from duration, assuming linear proportionality between moment and duration. It seems that the conclusion holds despite this probably incorrect scaling assumption.

Finally, we note that the zone of SSEs can be divided into a discrete number of segments that slip systematically as a whole, either independently or jointly (Fig. 2). From the rupture patterns, cumulative slip distribution and number of times that a sub-fault has slipped (Extended Data Fig. 3), we defined 13 segments (separated by dashed red lines in

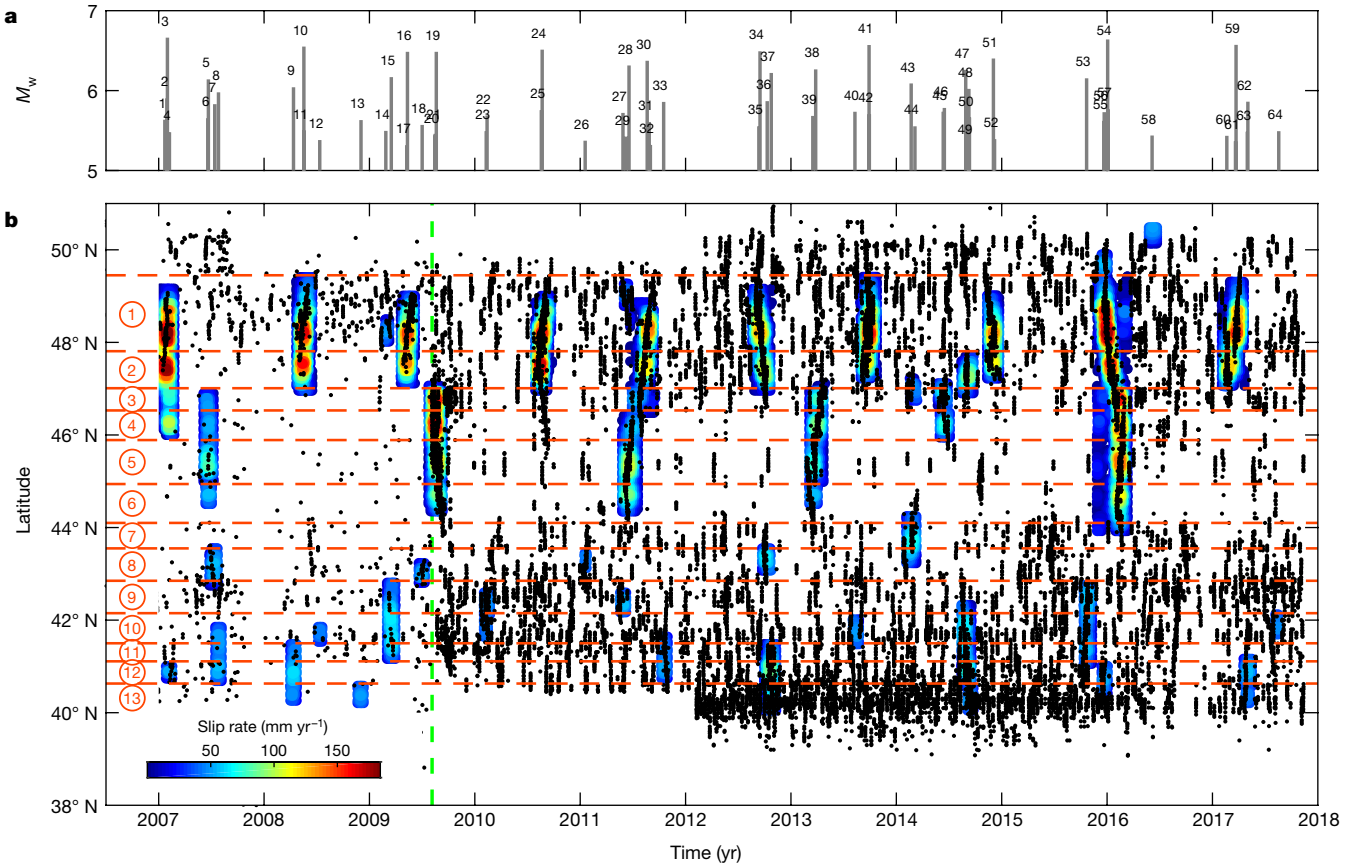



**Fig. 2 | Spatio-temporal distribution and segmentation of SSEs. a**, Timeline with magnitudes, labelled by event number, of all 64 SSEs of our original catalogue[15]. **b**, Timing and rupture extent of the SSEs. The black dots indicate tremors. The catalogue from Ide[33] is used until 2009.595, and the catalogue from PNSN (https://pnsn.org/tremor) is used thereafter. The vertical green dashed line marks the separation between the two catalogues. The dashed red lines indicate the segment boundaries defined as the rupture ends (see also Extended Data Fig. 3).

Fig. 2b). Segments 1 and 2 are extremely strongly coupled. They mostly rupture together except for a rupture in July 2014 (2014.612) that was restricted to segment 2. Segment 7 ruptured in combination with segments 6 and 8 in 2014, but never by itself. The segmentation of the Cascadia SSE zone had already been noticed[7], and a similar segmentation is observed in Japan[29]. This segmentation is qualitatively similar to the segmentation defined by regular megathrust earthquakes[30,31].

In conclusion, the $M_0 \propto T$ scaling proposed by Ide et al.[6] probably arises from the assembling of SSEs occurring under different conditions. We suspect that, as described here for the particular case of the SSEs in Cascadia, any subset of SSEs under similar conditions would yield a cubic scaling law as we found here. The along-strike segmentation, frequency–magnitude distribution and scaling properties of SSEs on the Cascadia subduction zone are thus remarkably similar to those of regular earthquakes. The pulse-like propagation of individual events also looks very similar to the seismic ruptures as inferred for large SSEs in the context of the Mexican subduction[32]. We infer that the dynamics governing aseismic SSEs is not very different from the dynamics governing seismic ruptures, a surprising result given that seismic ruptures are commonly thought to be governed by inertial effects, which are needed to account for the radiated seismic waves, but should not have any role for SSEs. Unexpectedly, our results call for re-examination of the cause of the $M_0 \propto A^{3/2}$ scaling, as it seems that, at least in the case of SSEs, the explanation based on self-similarity and the circular crack model will not hold. A re-examination of the effect of geometric bounding on scaling properties of regular earthquakes as well as SSEs may be needed. The similar scaling properties of SSEs and regular earthquakes suggest that SSEs might be useful in developing and testing dynamic models of earthquake sequences that are difficult to constrain from observations of regular earthquakes, especially given the long return period of large earthquakes.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1673-6.


1. Dragert, H., Wang, K. & James, S. T. A silent slip event on the deeper Cascadia subduction interface. *Science* **292**, 1525–1528 (2001).
2. Ozawa, S. et al. Detection and monitoring of ongoing aseismic slip in the Tokai region, Central Japan. *Science* **298**, 1009–1012 (2002).
3. Lowry, A. R., Larson, K. M., Kostoglodov, V. & Bilham, R. Transient fault slip in Guerrero, southern Mexico. *Geophys. Res. Lett.* **28**, 3753–3756 (2001).
4. Douglas, A., Beavan, J., Wallace, L. & Townend, J. Slow slip on the northern Hikurangi subduction interface, New Zealand. *Geophys. Res. Lett.* **32**, L16305 (2005).
5. Bürgmann, R. The geophysics, geology and mechanics of slow fault slip. *Earth Planet. Sci. Lett.* **495**, 112–134 (2018).
6. Ide, S., Beroza, G. C., Shelly, D. R. & Uchide, T. A scaling law for slow earthquakes. *Nature* **447**, 76–79 (2007).
7. Gao, H. Y., Schmidt, D. A. & Weldon, R. J. Scaling relationships of source parameters for slow slip events. *Bull. Seismol. Soc. Am.* **102**, 352–360 (2012).
8. Peng, Z. G. & Gomberg, J. An integrated perspective of the continuum between earthquakes and slow-slip phenomena. *Nat. Geosci.* **3**, 599–607 (2010).
9. Kanamori, H. & Anderson, L. Theoritical basis of some empirical relations in seismology. *Bull. Seismol. Soc. Am.* **65**, 1073–1095 (1975).
10. Gomberg, J., Wech, A., Creager, K., Obara, K. & Agnew, D. Reconsidering earthquake scaling. *Geophys. Res. Lett.* **43**, 6243–6251 (2016).

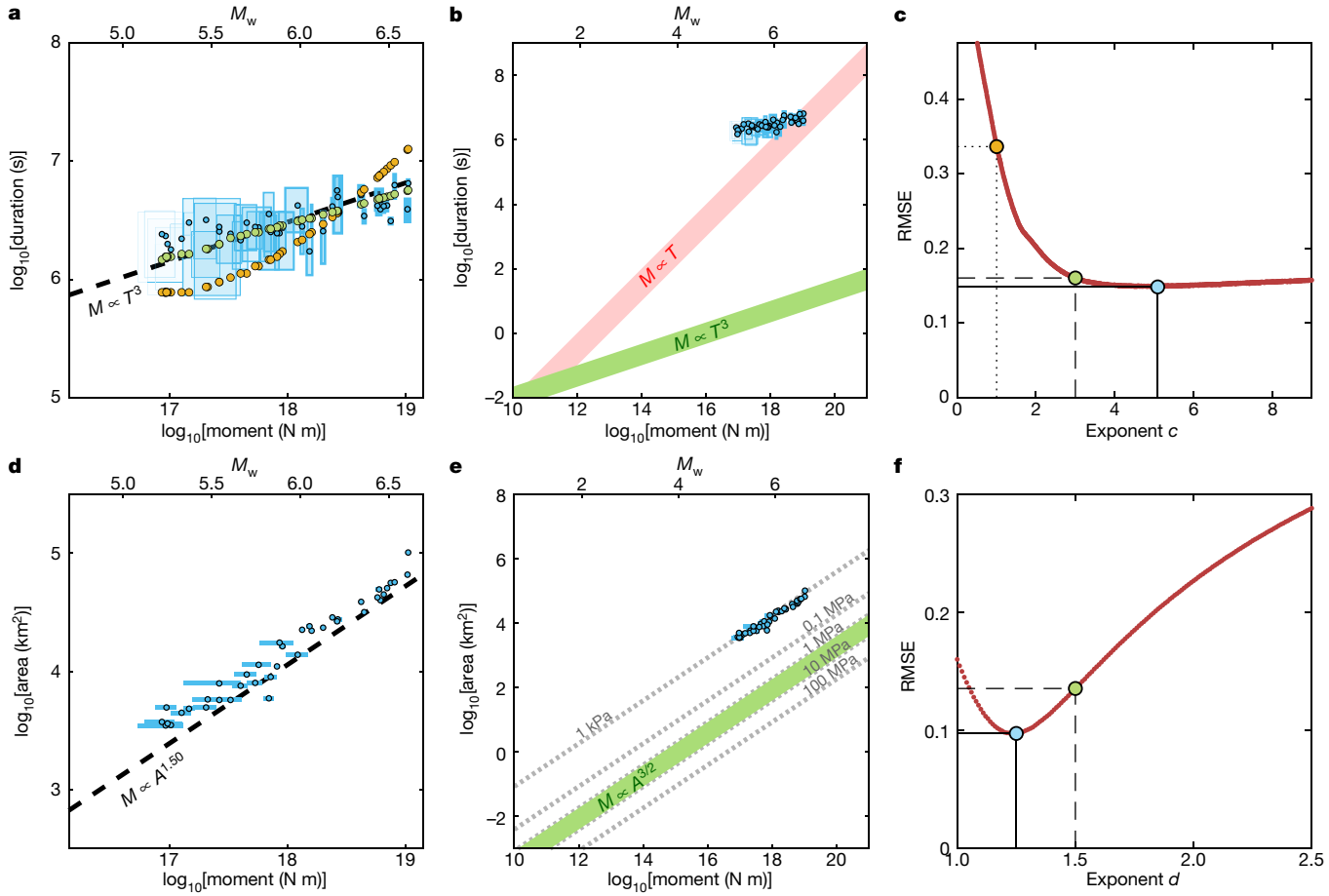



**Fig. 3 | Moment–duration and moment–area scaling. a**, Relationship between moment and duration of SSEs, compared with cubic scaling of regular earthquakes (dashed line). The blue shaded boxes outline the domain defined by the minimum and maximum estimates of SSE duration and moment release. The blue filled circles represent the mean value between the minimum and maximum duration and moment estimations. Green and orange filled circles show expected moment–duration distribution for $M_0 \propto T^3$ and $M_0 \propto T$, after accounting for the temporal filter effect. **b**, Comparison with proposed[6] scaling for slow (red) and regular earthquakes (green). **c**, Data fit assuming $M_0 \propto T^c$, taking into account magnitude and duration uncertainties and temporal filter effect (see Methods). The RMSE for $c = 3$ (green dot) is half that of $c = 1$ (yellow dot), and only 10% larger than the best fitting value $c = 5.09$ (blue dot). **d**, Relationship between moment and rupture area of SSEs compared to the 3/2 scaling of regular earthquakes. **e**, Comparison with regular earthquakes (green shading). Stress drop isolines (dotted grey lines) are based on the circular crack model[9]. **f**, Data fit assuming $M_0 \propto A^d$, taking into account the moment uncertainty (see Methods). The RMSEs for $d = 1.25$ (best value) and $d = 1.5$ are indicated by the green and blue dot, respectively.

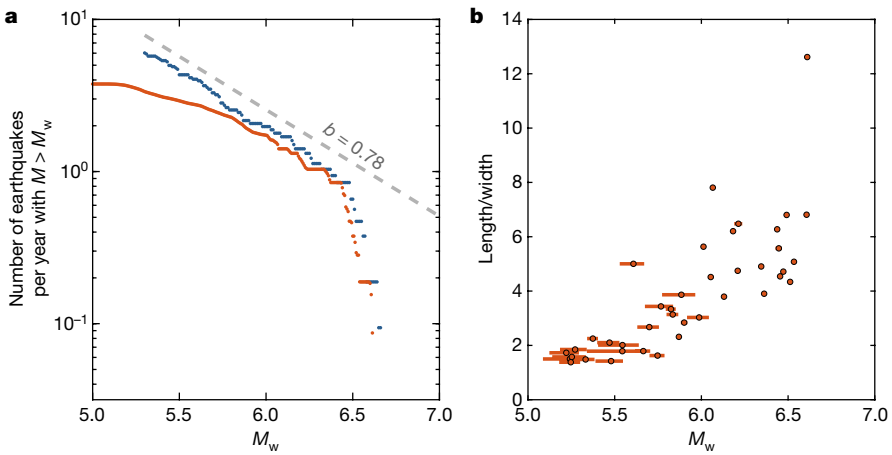



**Fig. 4 | Frequency–magnitude distribution and aspect ratio of SSEs in Cascadia. a**, Logarithm of the number of SSEs with moment magnitude larger than the value in abscissa using the original catalogue[15] (blue filled circles) and the revised catalogue (orange filled circles). Like regular earthquakes, SSEs are observed to follow approximately a linear trend, that is, the Gutenberg–Richter relationship (see Methods). The apparent larger $b$ value at $M_w > 6.4$, defined by only 11 events, could suggest that the distribution is truncated possibly as result of the along-strike segmentation. **b**, Aspect ratio of rupture areas. See Supplementary Information for details about measurements of area and aspect ratio.

11. Ben-Zion, Y. H. Episodic tremor and slip on a frictional interface with critical zero weakening in elastic solid. *Geophys. J. Int*. **189**, 1159–1168 (2012).
12. Hawthorne, J. C. & Bartlow, N. M. Observing and modeling the spectrum of a slow slip event. *J. Geophys. Res. Solid Earth* **123**, 4243–4265 (2018).
13. Romanet, P., Bhat, H. S., Jolivet, R. & Madariaga, R. Fast and slow slip events emerge due to fault geometrical complexity. *Geophys. Res. Lett*. **45**, 4809–4819 (2018).
14. Ide, S. A Brownian walk model for slow earthquakes. *Geophys. Res. Lett*. **35**, L17301 (2008).
15. Michel, S., Gualandi, A. & Avouac, J.-P. Interseismic coupling and slow slip events on the Cascadia megathrust. *Pure Appl. Geophys*. **176**, 3867–3891 (2019).
16. Rogers, G. & Dragert, H. Episodic tremor and slip on the Cascadia subduction zone: the chatter of silent slip. *Science* **300**, 1942–1943 (2003).
17. Obara, K., Hirose, H., Yamamizu, F. & Kasahara, K. Episodic slow slip events accompanied by non-volcanic tremors in southwest Japan subduction zone. *Geophys. Res. Lett*. **31**, L23602 (2004).
18. Madariaga, R. in *Encyclopedia of Complexity and System Science* (ed. Meyers, R.) 2581–2600 (Springer, 2009).
19. Denolle, M. A. & Shearer, P. M. New perspectives on self-similarity for shallow thrust earthquakes. *J. Geophys. Res. Solid Earth* **121**, 6533–6565 (2016).
20. Bartlow, N. M., Miyazaki, S., Bradley, A. M. & Segall, P. Space–time correlation of slip and tremor during the 2009 Cascadia slow slip event. *Geophys. Res. Lett*. **38**, L18309 (2011).
21. Wech, A. G. & Bartlow, N. M. Slip rate and tremor genesis in Cascadia. *Geophys. Res. Lett*. **41**, 392–398 (2014).
22. Galetzka, J. et al. Slip pulse and resonance of the Kathmandu basin during the 2015 Gorkha earthquake, Nepal. *Science* **349**, 1091–1095 (2015).
23. Gao, X. & Wang, K. L. Rheological separation of the megathrust seismogenic zone and episodic tremor and slip. *Nature* **543**, 416–419 (2017).
24. Scholz, C. H. *The Mechanics of Earthquakes* (Cambridge Univ. Press, 1990).
25. Noda, H., Lapusta, N. & Kanamori, H. Comparison of average stress drop measures for ruptures with heterogeneous stress change and implications for earthquake physics. *Geophys. J. Int*. **193**, 1691–1712 (2013).
26. Meade, B. J. Algorithms for the calculation of exact displacements, strains, and stresses for triangular dislocation elements in a uniform elastic half space. *Comput. Geosci*. **33**, 1064–1075 (2007).
27. Schmidt, D. A. & Gao, H. Source parameters and time-dependent slip distributions of slow slip events on the Cascadia subduction zone from 1998 to 2008. *J. Geophys. Res. Solid Earth* **115**, B00A18 (2010).
28. Wech, A. G., Creager, K. C., Houston, H. & Vidale, J. E. An earthquake-like magnitude-frequency distribution of slow slip in northern Cascadia. *Geophys. Res. Lett*. **37**, L22310 (2010).
29. Obara, K. Phenomenology of deep slow earthquake family in southwest Japan: spatiotemporal characteristics and segmentation. *J. Geophys. Res. Solid Earth* **115**, B00A25 (2010).
30. Kanamori, H. & McNally, K. C. Variable rupture mode of the subduction zone along the Ecuador–Colombia coast. *Bull. Seismol. Soc. Am*. **72**, 1241–1253 (1982).
31. Thatcher, W. Order and diversity in the modes of circum-Pacific earthquake recurrence. *J. Geophys. Res*. **95**, 2609–2623 (1990).
32. Radiguet, M. et al. Spatial and temporal evolution of a long term slow slip event: the 2006 Guerrero slow slip event. *Geophys. J. Int*. **184**, 816–828 (2011).
33. Ide, S. Variety and spatial heterogeneity of tectonic tremor worldwide. *J. Geophys. Res. Solid Earth* **117**, B03302 (2012).

# Methods

### Moment–duration scaling

We determine the best-fitting moment–duration scaling law, $\log_{10}(T) = \left(\frac{1}{c}\right)\log_{10}(M_0) + g$ where $M_0$ is expressed in N m and $T$ in seconds. We take into account the uncertainties in SSEs duration and moment (see Supplementary Information) and the effect of the temporal filter. We grid search for the best exponent, $c$, and intercept, $g$, of the scaling law. For each pair of exponent and intercept values, 1,500 random catalogues of 40 SSEs are created, assuming a uniform probability between the minimum and maximum values of moment and duration. We then compare these catalogues with the moments and durations of synthetic catalogues. The events in the synthetic catalogues have the same magnitudes as in the random catalogue, and thus the same final moment released, $M_s$, but a duration, $D_s$, prescribed by the tested scaling law. To account for the filter, we generate synthetic time series assuming boxcar moment rate function with a moment rate equal to $M_s/D_s$ during the event (and 0 N m d$^{-1}$ otherwise). We apply the same filter as to the real data (a zero-phase digital filtering, using a 5-day window), and estimate durations from the filtered moment rate functions (we take a moment rate threshold of $\dot{M}_{0\text{thresh}}$= 6.63 N m d$^{-1}$, equivalent to the case of the smallest subfault, as defined by the SSE kinematic model, slipping at 40 mm yr$^{-1}$ with a shear modulus $\mu$ = 30 GPa). Finally, for the tested exponent and intercept, the RMSE is calculated between the durations of the 40 × 1,500 events produced and their associated smoothed synthetics. The range of values explored for the intercept, $g$, and exponent, $c$, spans from −35 to 7, and 0.5 to 9, respectively, using a step of 0.01 for both. The minimum RMSE for each tested exponent is shown in Fig. 3c. The best fit corresponds to $c$ = 5.09 but is only about 8% smaller than the RMSE obtained for $c$ = 3.

### Moment–area scaling

We use a similar procedure to search for the best-fitting moment–area scaling law, $\log_{10}(A) = \left(\frac{1}{d}\right)\log_{10}(M_0) + r$, taking into account the uncertainties in SSE moments. We grid search for the best exponent, $d$, and intercept, $r$. The area is expressed in km$^2$ and the moment in N m. For each pair of tested exponent and intercept, 1,500 random catalogues of 40 SSEs are created, assuming a uniform probability distribution between the estimated minimum and maximum values of moments. For each of these catalogues, an associated synthetic catalogue is created with areas prescribed to follow the tested scaling law. For each tested exponent and intercept, a RMSE is then calculated between the areas of the 40 × 1,500 events produced and their associated synthetics. The tested values of the intercept, $r$, and exponent, $d$, range from −15 to −1.5 and from 1 to 2.5 , respectively, using a step equal to 0.01 for both. The minimum RMSE for each exponent tested is shown in Fig. 3f and the best fit corresponds to an exponent equal to 1.25.

### Measurement of SSE rupture area and aspect ratio

We define $\dot{\delta}_{\text{deficit}}$ as the deficit of slip rate with respect to the long-term slip rate. The onset of SSEs are defined by $\dot{\delta}_{\text{deficit}} < V_{\text{thresh}}$. We set the slip deficit rate threshold, $V_{\text{thresh}}$, to −40 mm yr$^{-1}$. The area of a SSE is defined as the sum of the areas of the sub-fault with $\dot{\delta}_{\text{deficit}} < V_{\text{thresh}}$ (ref. [15]). This threshold is applied to a 30-day filtered $\dot{\delta}_{\text{deficit}}$. We include the neighbouring sub-faults in the calculation of the slipping area. We thus estimate the lengths and widths of the SSEs relative to a mean strike line that approximately follows the curved geometry of the megathrust and runs through the middle of the cumulated slip distribution of SSEs (Extended Data Fig. 3). For each SSE, the rupture length is defined as the distance between the northern and southern intersections of the rupture's outline with the mean strike line. The width is defined as twice the mean distance between the rupture's outline and the mean strike line. Because some SSE ruptures are not centred over it or might not even cut it, we shift the mean strike line along dip for each SSE, forcing it to pass through its slipping area where the measured length is maximum.

Note that the SSEs spatio-temporal extension is sensitive to the inversion regularization, the temporal filter applied to $\delta_{\text{deficit}}$ and the chosen value for $V_{\text{thresh}}$.

### Determination of magnitude-frequency distribution

The magnitude-frequency distribution for the revised catalogue in Fig. 4a is calculated taking into account the uncertainties in SSE magnitudes calculated in Supplementary Information 'Measurements of SSE duration and moment release'. We assume that each event has a uniform distribution within its moment uncertainty and sum all of those distributions. The resulting probability density function, $P_{\text{events}}$, gives the number of events per magnitude. We then calculate for each magnitude tested, $M_{\text{test}}$, the number of events greater than $M_{\text{test}}$ per year:

$$N = \int_{M_{\text{test}}}^{\infty} P_{\text{events}}(M_{\text{w}})\mathrm{d}M_{\text{w}}$$

The $b$ value of the Gutenberg–Richter distribution that best fits the original catalogue (64 events) is estimated to be 0.78 by using the maximum likelihood method[34]. We do not estimate the $b$ value for the revised catalogue owing to the rollover at lower magnitudes due to the data selection.

## Data availability

The durations and moments estimated in this study are listed in Extended Data Table 1 and in the Source Data of Fig. 3. The slip model of Michel et al.[15], which is used as input in this study is available at: ftp://ftp.gps.caltech.edu/pub/avouac/Cascadia_SSE_Nature/Data_for_Nature/.

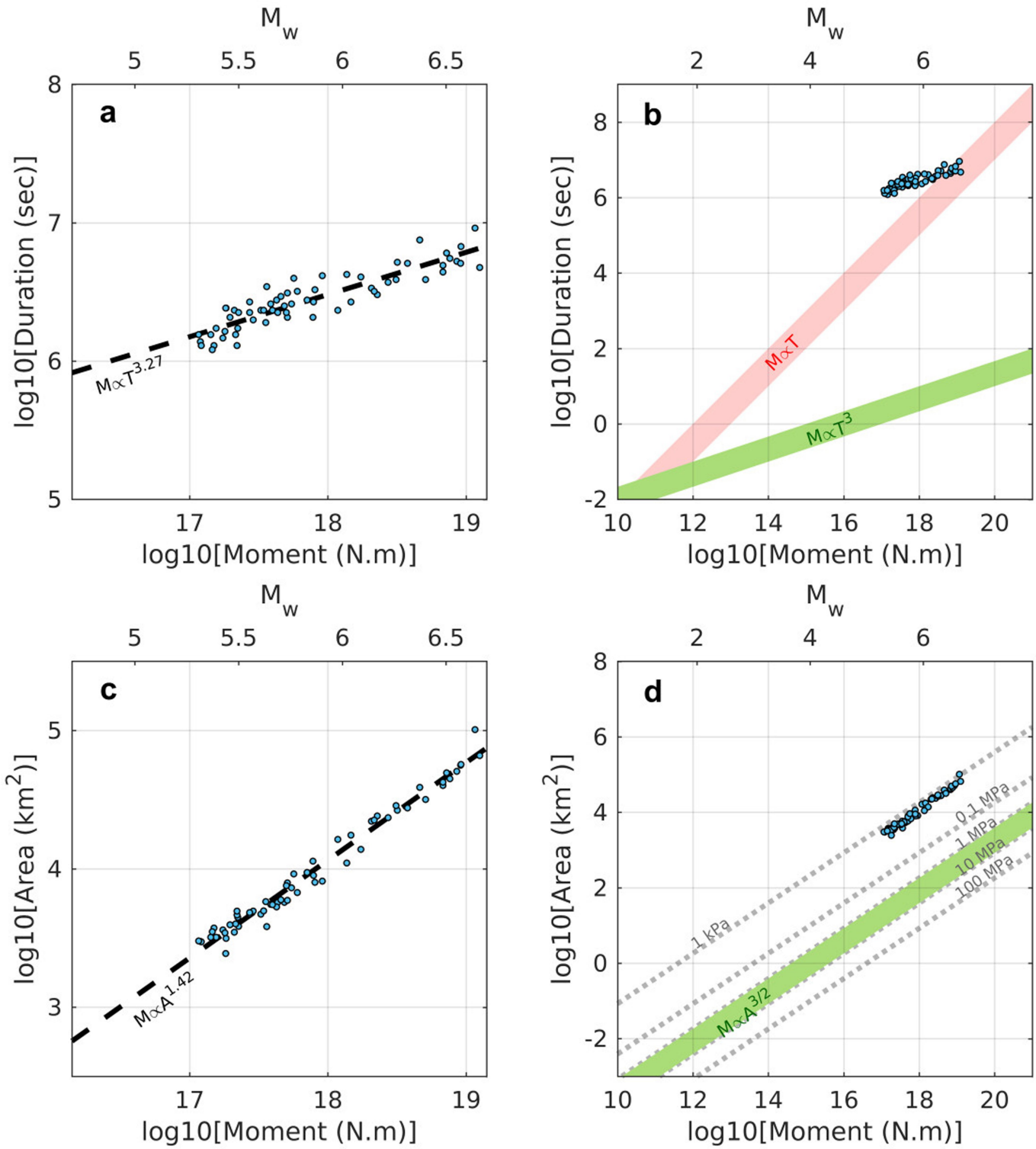
34. Aki, K. Maximum likelihood estimate of $b$ in the formula log $N = a - bM$ and its confidence limits. *Bull. Earthquake Res. Inst.* **43**, 237–239 (1965).

**Acknowledgements** This study was funded by NSF award EAR-1821853. S.M. is currently supported by a postdoctoral fellowship from CNES. We thank J. Gomberg for discussion and for providing a revised version of the catalogue of tremor durations presented in ref. [10]. We thank R. Burgmann for comments that helped to improve the study.

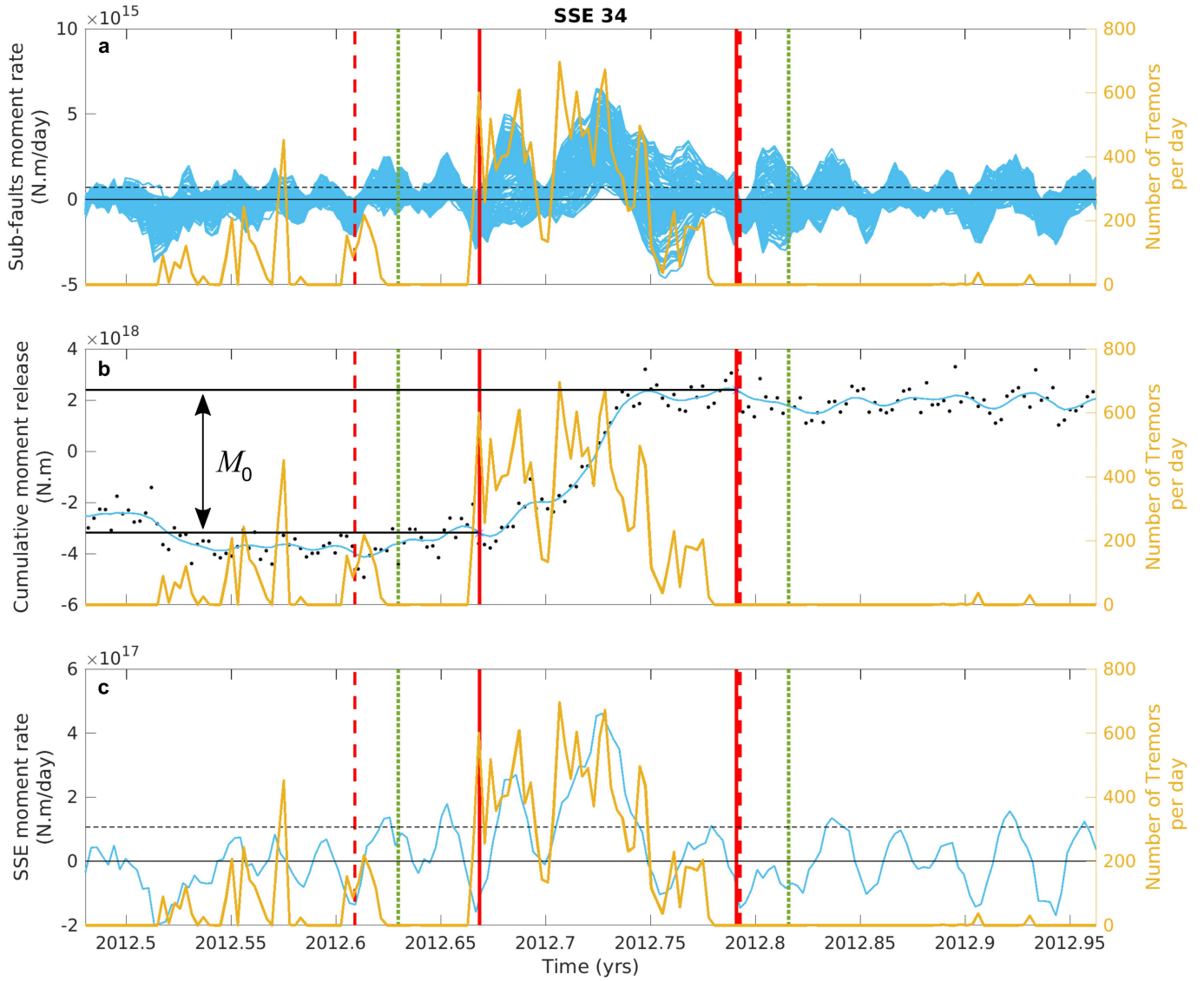
**Author contributions** S.M., A.G. and J.-P.A. designed the study, interpreted the results and wrote the manuscript; S.M. and A.G. performed the computations. J.-P.A. defined the scope of the study.

**Competing interests** The authors declare no competing interests.

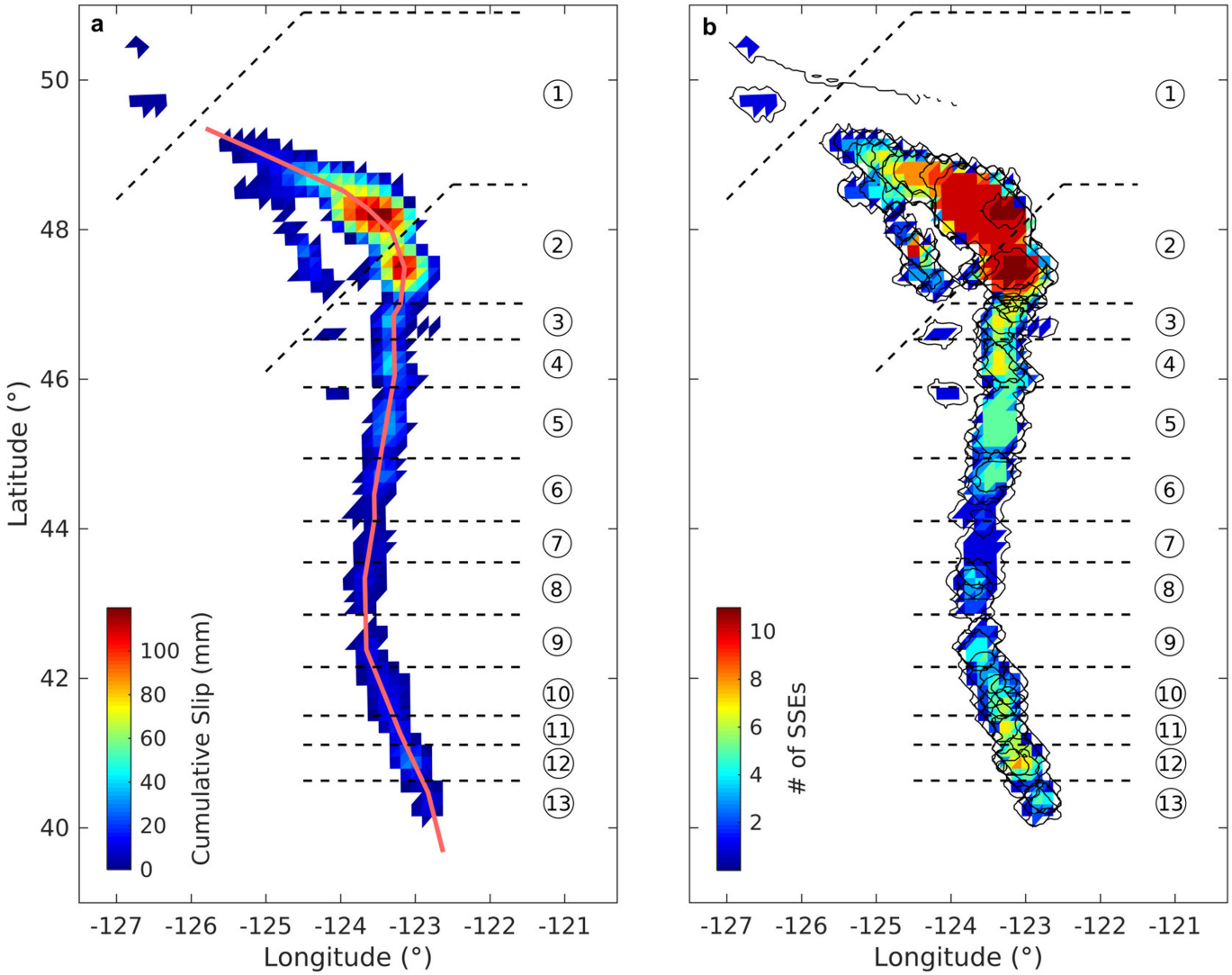
**Additional information**
**Supplementary information** is available for this paper at https://doi.org/10.1038/s41586-019-1673-6.
**Peer review information** *Nature* thanks Roland Burgmann, Ken Creager and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.
**Reprints and permissions information** is available at http://www.nature.com/reprints.



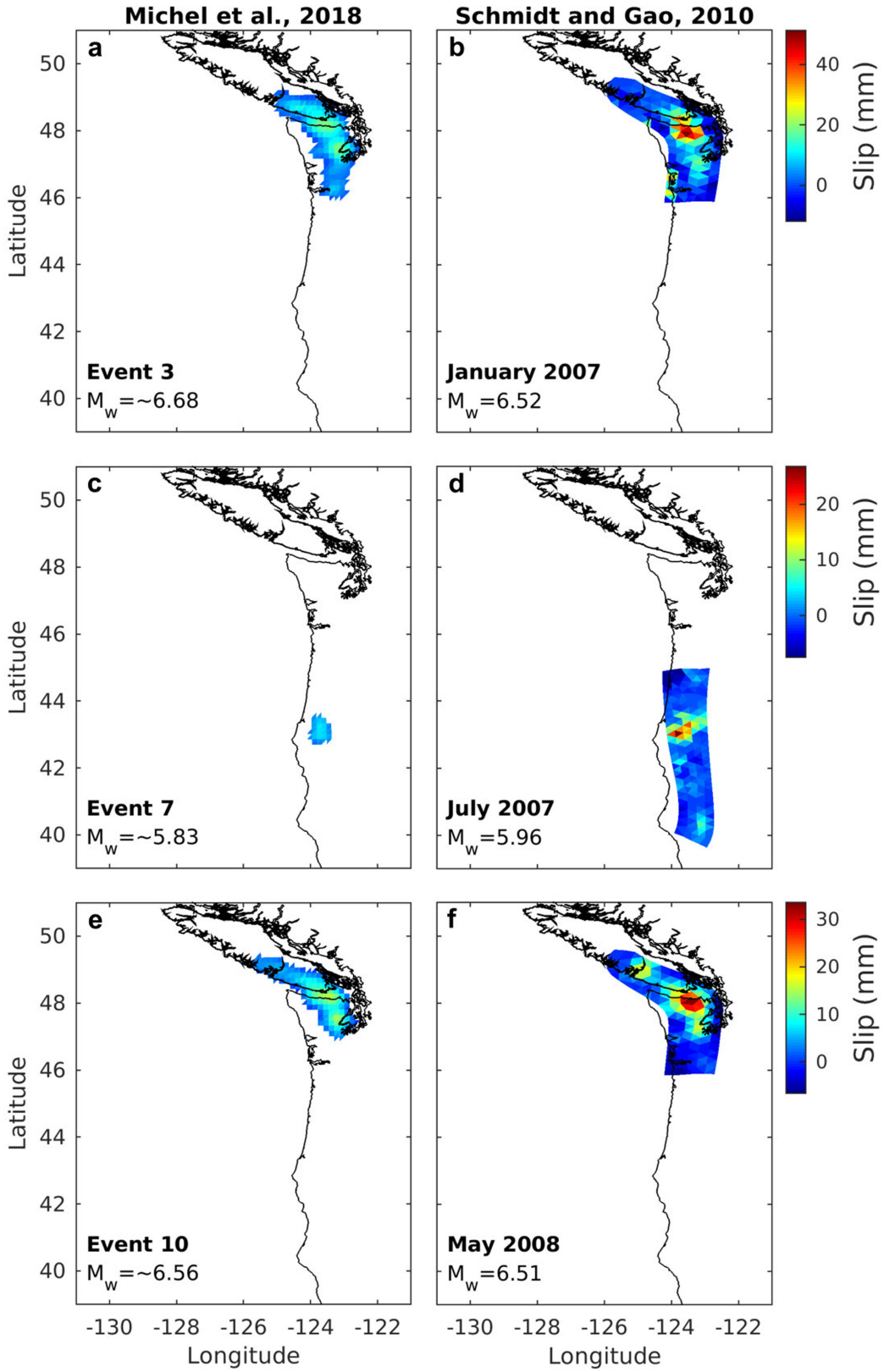

**Extended Data Fig. 1 | Moment–duration and moment–area scaling laws for automatic measurements. a**, Relationship between the moment released by SSEs and their duration. The black dashed line shows the best linear fit. **b**, Comparison with the scaling laws for slow (red shading) and regular (green shading) earthquakes proposed by Ide et al.[6]. **c**, Relationship between the moment released by SSEs and their rupture area. The black dashed line shows the best linear fit. **d**, Comparison with the scaling laws of regular earthquakes (green shading). Stress-drop isolines are estimated from the circular crack model.




**Extended Data Fig. 2 | SSEs duration estimations for the example of SSE 34.** **a**, The blue lines represent the moment rate evolution of each sub-fault participating in the SSE 34. The sub-faults moment rate is estimated using a zero-phase digital filtering on the rough $\delta_{deficit}$ using a 5-day window (effectively 9 days). The yellow line shows the number of tremors per day within the SSE rupture area. The solid red lines indicate the start and end times picked manually to estimate the minimum duration. They are determined by the timing of the first and last sub-faults with $\dot{M}_{0deficit} < \dot{M}_{0thresh}$ (the threshold rate is represented by the horizontal black dashed line). The dashed red lines similarly indicate the start and end times of the SSEs picked to estimate the maximum duration. They are determined by the times of the first and last sub-faults when $\dot{M}_{0deficit} < 0$. The dotted green lines indicate the automatic time picks for the start and end of the SSEs[15]. **b**, The black dots show the cumulative moment release in excess of the moment release that would have accumulated at the interseismic rate (as the SSE are extracted from the time series corrected for long-term interseismic strain). The blue line is its smoothed version using the same filter as indicated in **a**. The red and green vertical lines and the yellow curve are the same as in **a**. To illustrate the methodology used to calculate the SSE moment release, $M_0$, we indicate the values taken for the calculation based on the minimum duration by two horizontal solid black lines. **c**, The blue line indicates the SSE moment rate (sum of the moment rates of the SSE sub-faults). The horizontal black dashed line represents the sum of $\dot{M}_{0thresh}$ of all of the sub-faults. The red, yellow and green lines are the same as in **a**.




**Extended Data Fig. 3 | Segment delimitation. a**, SSEs cumulative slip. The pink line indicates a representative line of the average along-strike location of SSEs given by Michel et al.[15]. **b**, Map indicating the number of times that a sub-fault has experienced an SSE. The black contours delimit the extent of each SSE. The dashed black lines in **a** and **b** correspond to the selection of segments.




**Extended Data Fig. 4 | Comparison with slip models of a previously published study. a**, **c**, **e**, The cumulative slip models for SSEs 3, 7 and 10 of Michel et al.[15]. **b**, **d**, **f**, The cumulative slip models of the same SSEs estimated by Schmidt and Gao[27]. The magnitudes indicated in all panels are calculated by taking a shear modulus $\mu = 30$ GPa.

## Extended Data Table 1 | Manual estimation of SSE duration

| SSE # | Start (Max Duration) | Start (Min Duration) | End (Min Duration) | End (Max Duration) |
|---|---|---|---|---|
| 3 | 2007.0267 | 2007.0294 | 2007.128 | 2007.1773 |
| 4 | 2007.0294 | 2007.0733 | 2007.1034 | 2007.1472 |
| 5 & 6 | 2007.422 | 2007.4264 | 2007.5058 | 2007.5579 |
| 7 | 2007.4576 | 2007.4839 | 2007.5397 | 2007.5934 |
| 8 | 2007.491 | 2007.5318 | 2007.5852 | 2007.6099 |
| 9 | 2008.2286 | 2008.2642 | 2008.3025 | 2008.3464 |
| 10 | 2008.316 | 2008.3265 | 2008.4477 | 2008.4627 |
| 12 | 2008.4969 | 2008.5298 | 2008.5626 | 2008.6174 |
| 13 | 2008.8939 | 2008.924 | 2008.9569 | 2008.9745 |
| 14 | 2009.1266 | 2009.166 | 2009.213 | 2009.2197 |
| 15 | 2009.1759 | 2009.179 | 2009.2135 | 2009.251 |
| 16 | 2009.3183 | 2009.3238 | 2009.436 | 2009.4552 |
| 18 | 2009.429 | 2009.485 | 2009.5325 | 2009.5839 |
| 19 | 2009.5579 | 2009.587 | 2009.6989 | 2009.7125 |
| 22 & 23 | 2010.067 | 2010.0921 | 2010.1355 | 2010.178 |
| 24 | 2010.5859 | 2010.5887 | 2010.7064 | 2010.7324 |
| 26 | 2010.9993 | 2011.0431 | 2011.069 | 2011.087 |
| 27 | 2011.347 | 2011.372 | 2011.3936 | 2011.4867 |
| 28 | 2011.3689 | 2011.425 | 2011.5031 | 2011.6071 |
| 29 | 2011.3717 | 2011.4275 | 2011.451 | 2011.4747 |
| 30 | 2011.5305 | 2011.555 | 2011.6865 | 2011.7276 |
| 33 | 2011.7345 | 2011.796 | 2011.841 | 2011.864 |
| 34 | 2012.609 | 2012.6684 | 2012.791 | 2012.7926 |
| 36 | 2012.7242 | 2012.7269 | 2012.7844 | 2012.843 |
| 37 | 2012.7445 | 2012.7998 | 2012.8429 | 2012.8638 |
| 38 & 39 | 2013.1403 | 2013.1814 | 2013.305 | 2013.3758 |
| 40 | 2013.5401 | 2013.562 | 2013.5852 | 2013.6934 |
| 41 | 2013.6769 | 2013.682 | 2013.776 | 2013.781 |
| 43 | 2014.0274 | 2014.1314 | 2014.2108 | 2014.216 |
| 44 | 2014.119 | 2014.1218 | 2014.1971 | 2014.2357 |
| 45 & 46 | 2014.333 | 2014.438 | 2014.4928 | 2014.5051 |
| 47 & 50 | 2014.5914 | 2014.6051 | 2014.7135 | 2014.746 |
| 48 | 2014.6215 | 2014.6516 | 2014.69 | 2014.7392 |
| 51 | 2014.857 | 2014.8597 | 2014.955 | 2014.9802 |
| 53 | 2015.7276 | 2015.7851 | 2015.8371 | 2015.8535 |
| 54 & 55 | 2015.9521 | 2015.974 | 2016.168 | 2016.1711 |
| 56 | 2015.9202 | 2015.9603 | 2015.999 | 2016.0178 |
| 59 | 2017.039 | 2017.1239 | 2017.279 | 2017.2827 |
| 62 & 63 | 2017.2981 | 2017.2991 | 2017.3484 | 2017.4086 |
| 64 | 2017.5428 | 2017.5715 | 2017.603 | 2017.606 |

The start and end time pick for the minimum duration estimation are determined by the timing of the first and last sub-fault with $\dot{\delta}_{\text{deficit}} < V_{\text{thresh}}$. The start and end time picks for the maximum duration estimation are determined by the timing of the first and last sub-fault for which $\dot{\delta}_{\text{deficit}} < 0$. The SSE durations reported here are affected by the bias from the filter cut-off of about 9 days (see Supplementary Information 'Measurements of SSE duration and moment release').

Article

# Morphology of the earliest reconstructable tetrapod *Parmastega aelidae*

https://doi.org/10.1038/s41586-019-1636-y

Received: 8 April 2019

Accepted: 10 September 2019

Published online: 23 October 2019

Pavel A. Beznosov[1], Jennifer A. Clack[2], Ervīns Lukševičs[3], Marcello Ruta[4] & Per Erik Ahlberg[5]*

The known diversity of tetrapods of the Devonian period has increased markedly in recent decades, but their fossil record consists mostly of tantalizing fragments[1–15]. The framework for interpreting the morphology and palaeobiology of Devonian tetrapods is dominated by the near complete fossils of *Ichthyostega* and *Acanthostega*; the less complete, but partly reconstructable, *Ventastega* and *Tulerpeton* have supporting roles[2,4,16–34]. All four of these genera date to the late Famennian age (about 365–359 million years ago)—they are 10 million years younger than the earliest known tetrapod fragments[5,10], and nearly 30 million years younger than the oldest known tetrapod footprints[35]. Here we describe *Parmastega aelidae* gen. et sp. nov., a tetrapod from Russia dated to the earliest Famennian age (about 372 million years ago), represented by three-dimensional material that enables the reconstruction of the skull and shoulder girdle. The raised orbits, lateral line canals and weakly ossified postcranial skeleton of *P. aelidae* suggest a largely aquatic, surface-cruising animal. In Bayesian and parsimony-based phylogenetic analyses, the majority of trees place *Parmastega* as a sister group to all other tetrapods.

The rate of discovery of Devonian tetrapods accelerated greatly during the late twentieth and early twenty-first centuries. The description of *Ichthyostega* in 1932 was followed by *Acanthostega* in 1952, *Metaxygnathus* in 1977 and *Tulerpeton* in 1984; all other descriptions or identifications of genera (*Hynerpeton*, *Ventastega*, *Elginerpeton*, *Obruchevichthys*, *Densignathus*, *Sinostega*, *Jakubsonia*, *Ymeria*, *Webererpeton*, *Tutusius* and *Umzantsia*) as Devonian tetrapods have occurred since 1994[1–11,16,17]. Unnamed Devonian tetrapod material has previously been described from Belgium[12,13] and the USA[14,15]. However, the fossils of *Ichthyostega* and *Acanthostega* from East Greenland[9,16–31] remain by far the most complete material for Devonian tetrapods, followed by *Ventastega* fossils from Latvia[4,31,32] and *Tulerpeton* fossils from Russia[2,33,34]. All of these date to the final stage of the Devonian period (the late Famennian), by which point tetrapods had been in existence for about 30 million years (judging by evidence from trackways[35,36]) and had colonized both equatorial and polar environments[11]. Substantial differences between these four genera hint at long, divergent evolutionary histories; notably, the *Ichthyostega* and *Acanthostega* fossils have braincases that are fundamentally dissimilar to each other[20].

The tetrapod material described here is securely dated to the earliest Famennian age, and is comparable to that of *Ventastega* in its degree of completeness. It is derived from the Sosnogorsk Formation of the southern part of Timan Ridge (Komi Republic, Russia)[37], which straddles the boundary between the Frasnian age (about 382–372 million years ago) and the Famennian age; vertebrate remains occur only in the part of the Sosnogorsk Formation that dates to the Famennian age (Extended Data Fig. 1). The material described here is thus only marginally younger than the oldest known tetrapods *Elginerpeton*, *Obruchevichthys* and *Webererpeton*, which are known only from fragmentary material[5,10].

The quality of the material described here, which consists of numerous isolated bones and some articulated skull regions, is excellent. Multiple examples of the same bone all show the same distinctive features (Extended Data Fig. 2), which indicates that only a single tetrapod species is present (Extended Data Fig. 3). These fossils give a detailed picture of an animal from the earliest part of the known body-fossil record of the tetrapods.

## Systematic palaeontology

Tetrapoda Jaekel, 1909

*Parmastega aelidae* gen. et sp. nov.

**Remarks.** The term Tetrapoda is used here in its traditional, apomorphy-based sense of limbed vertebrates.

**Etymology.** The generic name derives from *parma*, a word in the Komi language describing the landscape of hills covered by coniferous forest, typical for South Timan, and *stégi* (Greek) meaning roof, understood here as the skull roof; *aelidae* honours Aelida I. Popova (Syktyvkar State University) (1929–2011), who first inspired P.A.B.'s interest in the natural sciences.

**Holotype.** Institute of Geology, Komi Science Centre (IG KSC) 705/1, an articulated snout region (Fig. 1a–c).

**Referred material.** One hundred and six individual bones or bone assemblies (Supplementary Table 1).

[1]Institute of Geology, Komi Science Centre, Ural Branch of the Russian Academy of Sciences, Syktyvkar, Russia. [2]University Museum of Zoology, University of Cambridge, Cambridge, UK. [3]Department of Geology, University of Latvia, Riga, Latvia. [4]Joseph Banks Laboratories, School of Life Sciences, University of Lincoln, Lincoln, UK. [5]Department of Organismal Biology, Uppsala University, Uppsala, Sweden. *e-mail: per.ahlberg@ebc.uu.se

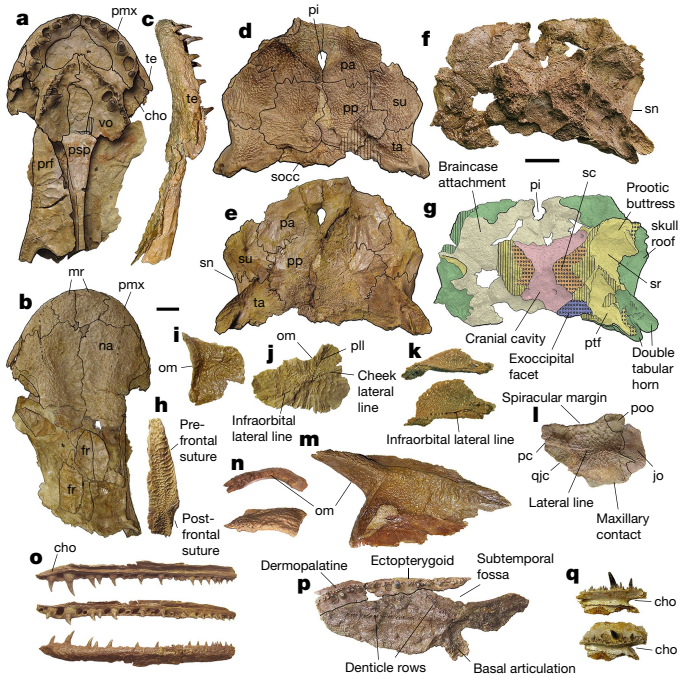



**Fig. 1 | Skull roof, cheek and palate of *P. aelidae*.** **a**–**c**, Specimen IG KSC 705/1 (all numerical codes in the figure legends refer to specimen numbers in the IG KSC collections), the holotype of *P. aelidae*. Articulated ethmosphenoid with associated prefrontal in ventral (**a**), dorsal (**b**) and lateral (**c**) views. **d**, **e**, 705/2. Skull table in dorsal (**d**) and ventral (**e**) views. **f**, **g**, 705/17. Skull table and partial braincase in ventral view. **g**, False-colour image identifying the components of 705/17. **h**, 705/18. Right frontal, dorsal view. **i**, 705/19. Left postorbital, external view. **j**, 705/20. Left jugal, external view. **k**, 705/25. Left lacrimal, lateral (top) and dorsal (bottom) views. **l**, 705/26. Right squamosal, external view. **m**, 705/5. Right prefrontal, external view. **n**, 705/4. Left postfrontal, lateral (top) and dorsal (bottom) views. **o**, 705/28. Right maxilla in internal (top), ventral (middle) and external (bottom) views. **p**, 705/29 (left dermopalatine), 705/30 (ectopterygoid) and 705/31 (pterygoid) in ventral view. **q**, 705/32. Left dermopalatine in lateral (top) and ventral (bottom) views. cho, choana; fr, frontal; jo, jugal overlap; mr, median rostrals; na, nasal; om, orbital margin; pa, parietal; pc, preopercular contact; pi, pineal foramen; pmx, premaxilla; pll, postorbital lateral line; poo, postorbital overlap; pp, postparietal; prf, prefrontal; psp, parasphenoid; ptf, posttemporal fossa; qjc, quadratojugal contact; sc, semicircular canals; socc, supraoccipital; sn, spiracular notch; sr, spiracular recess; su, supratemporal; ta, tabular; te, tectal; vo, vomer. Scale bars, 10 mm (scale bar below **f** applies to **f**, **g**; the scale bar below **c** applies to all other panels).

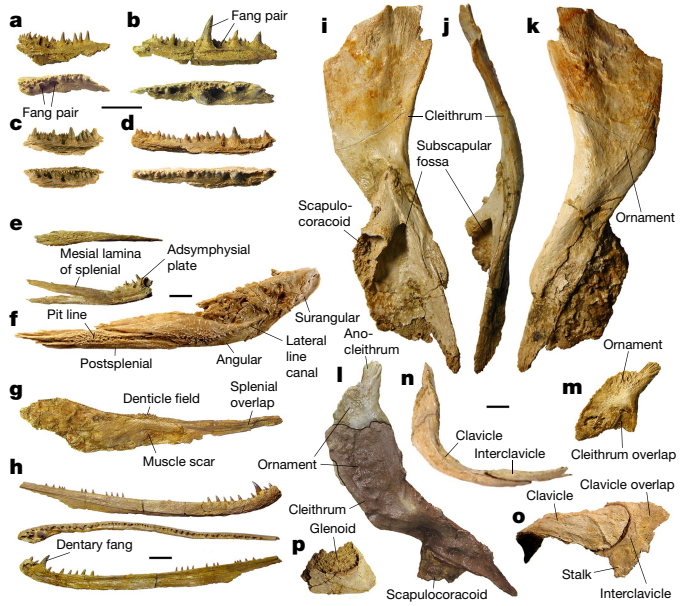



**Fig. 2 | Lower jaw and pectoral girdle of *P. aelidae*.** **a**, 705/21. Right adsymphysial plate in mesial (top) and dorsal (bottom) views. **b**, 705/22. Right anterior coronoid in mesial (top) and dorsal (bottom) views. **c**, 705/33. Right middle coronoid in mesial (top) and dorsal (bottom) views. **d**, 705/36. Left posterior coronoid in mesial (top) and dorsal (bottom) views. **e**, 705/37. Articulated left splenial and adsymphysial plate in ventrolateral (top) and mesial (bottom) views. **f**, 705/34. Articulated left postsplenial, angular and surangular in lateral view. **g**, 705/76. Left prearticular in mesial view. **h**, 705/67. Right dentary in lateral (top), dorsal (middle) and mesial (bottom) views. **i**–**k**, 705/15. Left cleithrum and partial scapulocoracoid in mesial (**i**), anterior (**j**) and lateral (**k**) views. **l**, 705/95 (right cleithrum) and 705/98 (anocleithrum) in lateral view. **m**, 705/98. Right anocleithrum in lateral view. **n**–**o**, 705/92 (right clavicle) and 705/89 (interclavicle) in anterior (**n**) and ventral (**o**) views. **p**, 705/102. Left coracoid in lateral view. Scale bars, 10 mm (scale bar between **a**–**d** applies to these four panels; **e**–**p** are all shown at the same scale).

**Locality and horizon.** Sosnovskiy Geological Monument, on the right bank of the Izhma River opposite Sosnogorsk (Komi Republic, Russia); Sosnogorsk Formation, lowermost Famennian age (Extended Data Fig. 1).

**Diagnosis.** A stem tetrapod diagnosed by the following unique combination of characters: dermal ornament of preorbital region developed into transverse parallel 'wave crests' with a spacing of a few millimetres; ornament present on dorsal blade of cleithrum and on anocleithrum; orbit strongly raised above skull roof, framed by an anterodorsal crest and a vertical anterior ridge carried on prefrontal; internasal fontanelle absent; median rostral paired; lacrimal excluded from orbit by prefrontal–jugal contact; intertemporal absent; pterygoids separated in midline by parasphenoid; interpterygoid vacuities absent; pterygoid dentition restricted to two lines of denticles, running anteriorly and anterolaterally from growth centre; ectopterygoid making large contribution to lateral wall of subtemporal fossa; middle part of otic capsule narrow, occupying approximately half of skull table width; posttemporal fossa wide, triangular; fang pair and row of marginal teeth on adsymphysial plate; middle part of prearticular with large muscle scar; interclavicle rounded with short posterior process.

## Description

The *Parmastega* material comprises the entire dermal skull (apart from the preopercular and the posterior part of the quadratojugal), the entire ethmoid and dorsal part of the otoccipital braincase, the entire lower jaw, the dermal pectoral girdle and the partly ossified scapulocoracoid (Figs. 1, 2). The material consists of a total of 106 numbered specimens (Supplementary Tables 1, 2), which represents a minimum of 11 individuals; these individuals show a wide range of sizes (Extended Data Figs. 2, 4), but were found within a small area of the site (Extended Data Fig. 1). Most specimens are isolated bones, but an articulated ethmoid (Fig. 1a–c) and several skull tables (Fig. 1d–g) are also present. The bones are three-dimensionally preserved in limestone with little or no distortion, and have been freed from the matrix using dilute acetic acid (Methods). Bones from the same individual can sometimes be identified by matching size and sutural fit (Extended Data Fig. 3). This allows us to reconstruct the skull, lower jaw and pectoral girdle with a high degree of confidence, except for the posterior part of the suspensorium (Fig. 3). Assuming proportions similar to those of *Acanthostega*[19], the maximum length of *Parmastega* was approximately 130 cm.

The shape of the skull is broadly similar to that of *Ventastega* and *Acanthostega*, although the orbits of *Parmastega* are raised higher above the skull table and the snout has a distinctly concave profile (Extended Data Fig. 4). The strongly raised orbits and relatively narrow snout are

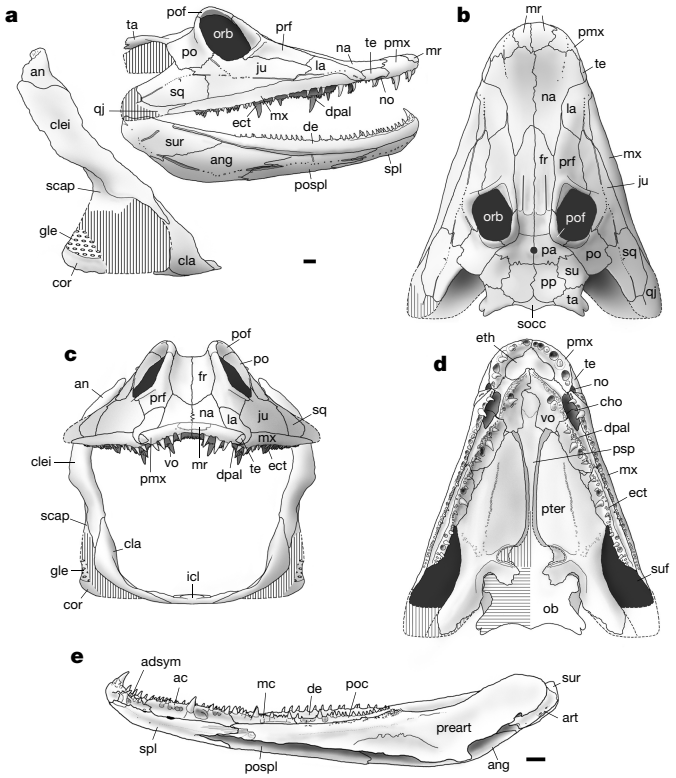



**Fig. 3 | Reconstructions of *P. aelidae*.** **a**, Skull, lower jaw and pectoral girdle of *Parmastega* in right lateral view. **b**, Skull in dorsal view. **c**, Skull and pectoral girdle in anterior view. **d**, Skull in ventral view. **e**, Right lower jaw ramus in mesial view. ac, anterior coronoid; adsym, adsymphysial plate; an, anocleithrum; ang, angular; art, articular; cla, clavicle; clei, cleithrum; cor, coracoid; de, dentary; dpal, dermopalatine; ect, ectopterygoid; eth, ethmoid; gle, glenoid; icl, interclavicle; ju, jugal; la, lacrimal; mc, middle coronoid; mx, maxilla; no, nostril; orb, orbit; ob, otoccipital braincase; po, postorbital; poc, posterior coronoid; pof, postfrontal; pospl, postsplenial; preart, prearticular; pter, pterygoid; qj, quadratojugal; scap, scapula; spl, splenial; sq, squamosal; suf, subtemporal fossa; sur, surangular. Vertical hatching indicates a missing element with unknown outline; horizontal hatching indicates a damaged object with known outline. Scale of reconstruction determined by largest individual. Scale bars, 10 mm (**a**–**d** are all shown to the same scale, which is given in **a**).

reminiscent of the elpistostegids *Elpistostege* and *Tiktaalik*[38,39]. However, the orbits of *Parmastega* are proportionately larger than those in the elpistostegids (Extended Data Fig. 5).

The dermal bone pattern of the skull roof and cheeks is, with a single exception, characteristic of Devonian tetrapods. There is no postrostral mosaic or internasal fontanelle. The median rostral is paired as in *Acanthostega*, *Ventastega* and *Elpistostege*, but unlike in *Ichthyostega* and *Elginerpeton* in which it is single[5,18,26,32,38]. A tectal bone forms the dorsal margin of the naris, which lies very close to the jaw margin and faces ventrally; the ventral margin of the naris is formed by the maxilla, as there is no lateral rostral. The lacrimal is excluded from the orbit by a long suture between the jugal and prefrontal. The latter is elongate and carries two bony crests, one forming the anterior part of the 'eyebrow' and the other an oblique ridge in front of the orbit; both are more strongly developed in large specimens (Figs. 1m, 3a–c). The frontals are elongate with a distinct transverse 'step' on the posterior part of the dorsal surface, marking the transition from snout to skull table. Intertemporals are absent. The lateral margins of the supratemporal and tabular form a raised spiracular margin; the tabular horn has distinct dorsal and ventral components. A small part of the dorsal surface of the braincase is exposed posterior to the tabulars. The dermal ornament of the preorbital region includes areas of irregular transverse ripples (Fig. 1h, m, Extended Data Fig. 2), somewhat similar to the ornament of *Umzantsia*[11] but much coarser; elsewhere, the ornament grades into the conventional tetrapod 'starburst' ornament. Partly enclosed sensory-line canals are well-developed on the premaxilla, cheek bones and the anterior part of the nasals, but are absent from the skull table (Fig. 1d).

Between the anterior suture for the jugal and the posterior suture for the preopercular, the ventral margin of the squamosal presents two distinct sutural margins that appear to be contacts for two bones (Fig. 1l). The posterior of these margins must be for the quadratojugal; as the jugal lacks a posterior process, we tentatively infer that the anterior segment of the ventral margin of the squamosal contacts the maxilla (Fig. 3a). A squamosal–maxillary contact is characteristic for 'fish' members of the tetrapod stem-group (such as *Eusthenopteron*[40]) and its presence in *Parmastega* is unique among tetrapods.

The palatal morphology of *Parmastega* is intermediate between that of the elpistostegids and that of Devonian tetrapods. In the elpistostegids *Panderichthys* and *Tiktaalik*, the pterygoids are separated in the midline by a long denticulated parasphenoid[41,42]. The vomer has a transverse posterior margin; in *Panderichthys*, this margin ends mesially in a short posterior process that extends along the lateral margin of the parasphenoid[41]. This condition is broadly similar to that observed in *Eusthenopteron*[40]. By contrast, in *Ichthyostega*, *Acanthostega* and *Ventastega*, the pterygoids meet in the midline (separating the parasphenoid from the vomers) and the most-posterior point of the vomer is its posterolateral corner[4,18,23]. In *Parmastega*, the parasphenoid separates the pterygoids but is not denticulated anteriorly, and the vomeral morphology is intermediate between that of *Panderichthys* and Devonian tetrapods (Figs. 1a, 3d). The pterygoid carries a longitudinal row or narrow band of denticles, and a shorter oblique band that extends anterolaterally. Uniquely, the ectopterygoid extends posteriorly past its contact with the pterygoid to contribute to the lateral margin of the subtemporal fossa (Fig. 3d). This relationship is demonstrated by a sutural fit of three bones from one individual (Fig. 1p).

Two parts of the braincase are preserved: the ethmoid and part of the sphenoid in IG KSC 705/1, and the dorsal part of the otoccipital in IG KSC 705/17 (Fig. 1a, f, g). An ossified ethmoid is shared only with *Ichthyostega* among known Devonian tetrapods[18]. The otoccipital has a strongly developed pro-otic buttress, a narrow cranial cavity with small inner ears and a posttemporal fossa that is bounded laterally by a crista parotica that extends onto the tabular horn. In ventral view, the outline of the occipital resembles that of *Tiktaalik*[42], but is proportionately broader in *Parmastega*. Otoccipitals that are previously known from Devonian tetrapods show one of two radically different morphologies. In *Acanthostega* and *Ventastega*, the narrow posttemporal fossa is open laterally and the braincase occupies almost the whole ventral surface of the skull table; by contrast, the narrow braincase in *Ichthyostega* is flanked by large cavities under the skull table that probably housed spiracular diverticula[20,24,25,32]. The otoccipital of *Parmastega* provides a plausible ancestral groundplan for both of these morphologies (Extended Data Fig. 6).

The construction of the lower jaw is typical for tetrapods[30], although it is unusually slender and delicate (Figs. 2a–h, 3e). The only ossified parts of the Meckelian element are the articular and the symphysis. The prearticular carries very few denticles but bears a large ventral muscle scar on its middle part. The contact between the prearticular and the mesial lamina of the splenial is not a tight suture, as in other known Devonian tetrapods[30], but is instead a loose overlap that must have contained a ligamentous component and allowed a degree of flexibility. Fang pairs—positioned mesial to the tooth row—are present on the adsymphysial plate, dentary, and anterior and middle coronoids. Postsplenial and surangular pit lines are present. The dentary is splint-like and loosely attached.

The pectoral girdle is U-shaped in anterior view, and the dorsal blades of the cleithra are approximately parallel (Figs. 2i–o, 3a, c). The dorsal orientation of the anocleithrum, which we determined from well-preserved

contact surfaces, makes the girdle notably tall. The cleithrum and anocleithrum both carry a dermal ornament, a characteristic that is otherwise absent in known tetrapods (except for *Umzantsia*[11]). The clavicle is narrow, and the interclavicle has a rounded corpus with a short posterior process (Fig. 2n, o); both of these bones somewhat resemble the corresponding elements in *Ichthyostega*[18], whereas *Acanthostega* and *Ventastega* have broader clavicles and lozenge-shaped interclavicles[29,32]. The scapulocoracoid is ossified in two parts: a dorsal scapular part that is co-ossified with the cleithrum (Fig. 2i) and a posterior coracoid ossification that carries the glenoid (Fig. 2p). As in *Ichthyostega*, *Elginerpeton* and *Hynerpeton*, the subscapular fossa is deep and has a narrow apex; by contrast, in *Acanthostega* and *Ventastega* the fossa is shallow and broad[3,18,29,32,43]. The limbs, pelvis, vertebrae and ribs are not preserved in the material from Sosnogorsk.

## Phylogenetic analysis

We evaluated the phylogenetic position of *Parmastega* with maximum parsimony and Bayesian inference analyses, applied to a data matrix of 26 taxa and 113 characters (Methods). The character list and data matrix are provided in the Supplementary Information.

The resolution of the strict-consensus, unweighted parsimony analysis was poor: all of the Devonian tetrapods (including *Parmastega*) formed a polytomy together with 'whatcheeriid-grade' Carboniferous taxa (Extended Data Fig. 7a). However, in 70% of the trees, *Parmastega* was the sister group to all other tetrapods. We used a range of approaches (character reweighting by rescaled consistency index and *K* values, and the calculation of agreement subtrees from consensus trees) to investigate the phylogenetic signal in the dataset (Extended Data Fig. 7b, c, e–h), which revealed consistent patterns. If the position of *Parmastega* was resolved, it was always placed as the sister group to all other tetrapods; if *Ventastega* was resolved, it was placed immediately crownward of *Parmastega*. *Ichthyostega* was resolved crownward of *Acanthostega* in the Adams consensus of unweighted trees, but in the reweighted analyses *Acanthostega* was crownward of *Ichthyostega*. The Bayesian tree (Extended Data Fig. 7d) also recovered these positions for *Parmastega* and *Ventastega*, but did not resolve *Ichthyostega* and *Acanthostega*.

## Discussion

*Parmastega* is morphologically intermediate between the elpistostegids *Tiktaalik*, *Elpistostege* and *Panderichthys* on the one hand, and previously known Devonian tetrapods on the other—but primitive and derived characters are not evenly distributed across its anatomy. The lower jaw, pectoral girdle, external dermal bone pattern of the snout region, the absence of gular plates and the relative size of the orbits are all tetrapod-like, whereas elpistostegid-like characteristics persist in the palate and the dermal ornamentation of the cleithrum and anocleithrum. Although no appendage bones are known, the morphology of the pectoral girdle strongly suggests that *Parmastega* possessed limbs rather than paired fins. The scapulocoracoid, which forms the proximal attachment for many forelimb muscles and undergoes substantial changes in shape from elpistostegids[44,45] to tetrapods[3,18,29,32,34], is particularly informative in this regard: *Parmastega* conforms to the tetrapod pattern. The shape and construction of the lower jaw, and the absence of gular plates, suggest that gill ventilation and prey capture worked in the same way as in more-crownward Devonian tetrapods. The reconfiguration of the palate and the loss of dermal ornament on the shoulder girdle evidently lagged behind these transformations.

Until now, one of the most puzzling aspects of the anatomy of Devonian tetrapods has been the specialized ear region of *Ichthyostega*, which differs markedly from the ear regions in other early tetrapods[18,20]. The braincase of *Parmastega* is morphologically intermediate between that of *Ichthyostega* and those of *Acanthostega* and *Ventastega*, providing a plausible hypothetical ancestor for both patterns (Extended Data Fig. 6a). However, these transformations cannot be mapped parsimoniously onto the phylogeny, indicating the presence of non-trivial homoplasy either in the braincases or in other parts of the skeleton (Extended Data Fig. 6b).

The three-dimensional preservation and apparent absence of post-mortem transport makes the *Parmastega* fossils palaeobiologically informative. The environment of preservation, which was probably also the living environment of *Parmastega*, was a coastal lagoon with brackish water and a rich fish fauna including the placoderm *Bothriolepis* and various sarcopterygians[46]. The concentration of the tetrapod remains in a small area of the site (Extended Data Fig. 1) suggests that *Parmastega* may have been a schooling animal. The vertebrate-bearing bed (bed 40, the 'fish dolomite') is composed of two consecutive tempestites; possibly a school of *Parmastega* was killed by the first storm event and their skeletons partly disarticulated by the second. Schooling behaviour is also implied by the mass occurrence of *Acanthostega* on Stensiö Bjerg (East Greenland)[47].

Raised orbits and a lack of lateral line canals on the skull table in *Parmastega* (Fig. 3a, b) suggests it adopted a surface-skimming position in the water, with emergent eyes, similar to that of extant crocodilians (Extended Data Fig. 8). The increase in orbit size across the transition between fish and tetrapods has previously been linked to a shift from aquatic to aerial vision[48]; the relative orbit size of *Parmastega* falls well within the tetrapod range (Extended Data Fig. 5) and its eyes were thus probably adapted for use in air. Although all known Devonian tetrapods have dorsally positioned eyes, *Parmastega* shows the most extreme condition (Extended Data Fig. 4). The nostrils of *Parmastega* face ventrally, which suggests that the nose was not used for breathing air (Extended Data Fig. 8). The dorsally placed spiracles may have taken on this function, as has previously been argued for *Panderichthys*[49] and more-crownward Devonian tetrapods[20,22]. Similar to the condition in *Ventastega*, *Acanthostega*[31] and *Ichthyostega*[18], the lower jaw does not match the upper jaw in curvature in lateral or in ventral view (Extended Data Fig. 9).

The *Parmastega* material contains no vertebrae, ribs, pelvic girdles or limb bones. The lack of evidence for post-mortem transport, the partially ossified nature of the scapulocoracoid even in the largest individuals and the preservation of the delicate isolated coracoid ossifications (Fig. 2i–l, p) suggest that this absence is not a taphonomic artefact but that it instead reflects a very lightly ossified, or even cartilaginous, axial and appendicular skeleton. *Ventastega* may also have had a lightly ossified postcranial skeleton[32]. *Acanthostega* and *Ichthyostega* became fully ossified as adults[17–19,21,27,29], but *Acanthostega* appears to have had a long juvenile stage with non-ossified endoskeleton[47]. Functionally, the poor ossification of *Parmastega* suggests little or no capacity for terrestrial locomotion. This contrasts strangely with the cranial morphology, which suggests that the eyes were habitually held above the surface of the water—and thus implies some kind of engagement with the terrestrial environment. Even more puzzling is the fact that this poorly ossified postcranial skeleton is apomorphic: elpistostegids are well-ossified, as are the majority of tetrapodomorph fishes[39,40,45].

*Parmastega* gives us the earliest detailed glimpse of a tetrapod: an aquatic, surface-skimming predator, just over a metre in length, living in a lagoon on a tropical coastal plain. *Parmastega* is phylogenetically the least-crownward of all of the non-fragmentary tetrapods, but it is not necessarily representative of the primitive conditions for the group. The slightly earlier *Elginerpeton*—which was also probably aquatic and was even larger than *Parmastega* (Extended Data Fig. 4)—had well-ossified girdles and limb bones, as well as a distinctive head shape with a narrow snout[5,30,43]. Moreover, the trackway record shows that tetrapods originated at least 20 million years before *Parmastega*[35,36], and the very existence of the trackways—which implies weight-bearing limbs, even if the prints were made in water—points to these forms having well-ossified postcranial skeletons. Together with the evidence for considerable

morphological homoplasy among Devonian tetrapods, this hints at a tangled and still-unknown early history for limbed vertebrates.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1636-y.


1. Campbell, K. S. W. & Bell, M. W. A primitive amphibian from the Late Devonian of New South Wales. *Alcheringa* **1**, 369–381 (1977).
2. Lebedev, O. A. The first find of a Devonian tetrapod in USSR [in Russian]. *Doklady Acad. Nauk SSSR* **278**, 1470–1473 (1984).
3. Daeschler, E. B., Shubin, N. H., Thomson, K. S. & Amaral, W. W. A Devonian tetrapod from North America. *Science* **265**, 639–642 (1994).
4. Ahlberg, P. E., Lukševičs, E. & Lebedev, O. The first tetrapod finds from the Devonian (Upper Famennian) of Latvia. *Phil. Trans. R. Soc. Lond. B* **343**, 303–328 (1994).
5. Ahlberg, P. E. *Elginerpeton pancheni* and the earliest tetrapod clade. *Nature* **373**, 420–425 (1995).
6. Daeschler, E. B. Early tetrapod jaws from the Late Devonian of Pennsylvania, USA. *J. Paleontol.* **74**, 301–308 (2000).
7. Zhu, M., Ahlberg, P. E., Zhao, W. & Jia, L. First Devonian tetrapod from Asia. *Nature* **420**, 760–761 (2002).
8. Lebedev, O. A. A new tetrapod *Jakubsonia livnensis* from the Early Famennian (Devonian) of Russia and palaeoecological remarks on the Late Devonian tetrapod habitats. *Acta Univ. Latviensis* **679**, 79–98 (2004).
9. Clack, J. A., Ahlberg, P. E., Blom, H. & Finney, S. M. A new genus of Devonian tetrapod from North-East Greenland, with new information on the lower jaw of *Ichthyostega*. *Palaeontology* **55**, 73–86 (2012).
10. Clément, G. & Lebedev, O. A. Revision of the early tetrapod *Obruchevichthys* Vorobyeva, 1977 from the Frasnian (Upper Devonian) of the north-western East European Platform. *Paleontol. J.* **48**, 1082–1091 (2014).
11. Gess, R. & Ahlberg, P. E. A tetrapod fauna from within the Devonian Antarctic Circle. *Science* **360**, 1120–1124 (2018).
12. Clément, G. et al. Devonian tetrapod from western Europe. *Nature* **427**, 412–413 (2004).
13. Olive, S. et al. New discoveries of tetrapods (ichthyostegid-like and whatcheeriid-like) in the Famennian (Late Devonian) localities of Strud and Becco (Belgium). *Palaeontology* **59**, 827–840 (2016).
14. Shubin, N. H., Daeschler, E. B. & Coates, M. I. The early evolution of the tetrapod humerus. *Science* **304**, 90–93 (2004).
15. Daeschler, E. B., Clack, J. A. & Shubin, N. H. Late Devonian tetrapod remains from Red Hill, Pennsylvania, USA: how much diversity? *Acta Zool.* **90**, 306–317 (2009).
16. Säve-Söderbergh, G. Preliminary note on Devonian stegocephalians from East Greenland. *Medd. Gronl.* **94**, 1–107 (1932).
17. Jarvik, E. On the fish-like tail in the ichthyostegid stegocephalians with descriptions of a new stegocephalian and a new crossopterygian from the Upper Devonian of East Greenland. *Medd. Gronl.* **114**, 1–90 (1952).
18. Jarvik, E. *The Devonian Tetrapod* Ichthyostega (*Fossils & Strata no. 40*) (Scandinavian Univ. Press, 1996).
19. Ahlberg, P. E., Clack, J. A. & Blom, H. The axial skeleton of the Devonian tetrapod *Ichthyostega*. *Nature* **437**, 137–140 (2005).
20. Clack, J. A. et al. A uniquely specialized ear in a very early tetrapod. *Nature* **425**, 65–69 (2003).
21. Callier, V., Clack, J. A. & Ahlberg, P. E. Contrasting developmental trajectories in the earliest known tetrapod forelimbs. *Science* **324**, 364–367 (2009).
22. Clack, J. A. Discovery of the earliest-known tetrapod stapes. *Nature* **342**, 425–427 (1989).
23. Clack, J. A. *Acanthostega gunnari*, a Devonian tetrapod from Greenland: the snout, palate and ventral parts of the braincase. *Medd. Gronl. Geosci.* **31**, 1–24 (1994).
24. Clack, J. A. Earliest known tetrapod braincase and the evolution of the stapes and fenestra ovalis. *Nature* **369**, 392–394 (1994).
25. Clack, J. A. The neurocranium of *Acanthostega gunnari* Jarvik and the evolution of the otic region in tetrapods. *Zool. J. Linn. Soc.* **122**, 61–97 (1998).
26. Clack, J. A. A revised reconstruction of the dermal skull roof of *Acanthostega gunnari*, an early tetrapod from the Late Devonian. *Trans. R. Soc. Edinb. Earth Sci.* **93**, 163–165 (2002).
27. Coates, M. I. & Clack, J. A. Polydactyly in the earliest known tetrapod limbs. *Nature* **347**, 66–69 (1990).
28. Coates, M. I. & Clack, J. A. Fish-like gills and breathing in the earliest known tetrapod. *Nature* **352**, 234–236 (1991).
29. Coates, M. I. The Devonian tetrapod *Acanthostega gunnari* Jarvik: postcranial anatomy, basal tetrapod interrelationships and patterns of skeletal evolution. *Trans. R. Soc. Edinb. Earth Sci.* **87**, 363–421 (1996).
30. Ahlberg, P. E. & Clack, J. A. Lower jaws, lower tetrapods – a review based on the Devonian genus *Acanthostega*. *Trans. R. Soc. Edinb. Earth Sci.* **89**, 11–46 (1998).
31. Porro, L. B., Rayfield, E. J. & Clack, J. A. Descriptive anatomy and three-dimensional reconstruction of the skull of the early tetrapod *Acanthostega gunnari* Jarvik, 1952. *PLoS ONE* **10**, e0118882 (2015).
32. Ahlberg, P. E., Clack, J. A., Lukševičs, E., Blom, H. & Zupiņs, I. *Ventastega curonica* and the origin of tetrapod morphology. *Nature* **453**, 1199–1204 (2008).
33. Lebedev, O. A. & Clack, J. A. Upper Devonian tetrapods from Andreyevka, Tula Region, Russia. *Palaeontology* **36**, 721–734 (1993).
34. Lebedev, O. A. & Coates, M. I. The postcranial skeleton of the Devonian tetrapod *Tulerpeton curtum*. *Zool. J. Linn. Soc.* **114**, 307–348 (1995).
35. Niedźwiedzki, G., Szrek, P., Narkiewicz, K., Narkiewicz, M. & Ahlberg, P. E. Tetrapod trackways from the early Middle Devonian period of Poland. *Nature* **463**, 43–48 (2010).
36. Stössel, I., Williams, E. A. & Higgs, K. T. Ichnology and depositional environment of the Middle Devonian Valentia Island tetrapod trackways, south-west Ireland. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **462**, 16–40 (2016).
37. Beznosov, P. Sosnogorsk Formation – a new local stratigraphic unit of the Upper Devonian from South Timan [in Russian]. *Geologiya i mineral'nye resursy evropeyskogo servero-vostoka Rossii: materialy XV geologicheskogo s'yezda Respubliki Komi* **2**, 9–12 (2009).
38. Schultze, H.-P. & Arsenault, M. The panderichthyid fish *Elpistostege*: a close relative of tetrapods? *Palaeontology* **28**, 293–310 (1985).
39. Daeschler, E. B., Shubin, N. H. & Jenkins, F. A. Jr. A Devonian tetrapod-like fish and the evolution of the tetrapod body plan. *Nature* **440**, 757–763 (2006).
40. Jarvik, E. *Basic Structure and Evolution of Vertebrates* Vol. 1 (Academic, 1980).
41. Ahlberg, P. E., Clack, J. A. & Lukševičs, E. Rapid braincase evolution between *Panderichthys* and the earliest tetrapods. *Nature* **381**, 61–64 (1996).
42. Downs, J. P., Daeschler, E. B., Jenkins, F. A. Jr & Shubin, N. H. The cranial endoskeleton of *Tiktaalik roseae*. *Nature* **455**, 925–929 (2008).
43. Ahlberg, P. E. Postcranial stem tetrapod remains from the Devonian of Scat Craig, Morayshire, Scotland. *Zool. J. Linn. Soc.* **122**, 99–141 (1998).
44. Vorobyeva, E. I. The shoulder girdle of *Panderichthys rhombolepis* (Gross) (Crossopterygii); Upper Devonian; Latvia. *Geobios* **28**, 285–288 (1995).
45. Shubin, N. H., Daeschler, E. B. & Jenkins, F. A. Jr. The pectoral fin of *Tiktaalik roseae* and the origin of the tetrapod limb. *Nature* **440**, 764–771 (2006).
46. Lukševičs, E., Beznosov, P. & Stūris, V. A new assessment of the Late Devonian antiarchan fish *Bothriolepis leptocheira* from South Timan (Russia) and the biotic crisis near the Frasnian–Famennian boundary. *Acta Palaeontol. Pol.* **62**, 97–119 (2017).
47. Sanchez, S., Tafforeau, P., Clack, J. A. & Ahlberg, P. E. Life history of the stem tetrapod *Acanthostega* revealed by synchrotron microtomography. *Nature* **537**, 408–411 (2016).
48. MacIver, M. A., Schmitz, L., Mugan, U., Murphey, T. D. & Mobley, C. D. Massive increase in visual range preceded the origin of terrestrial vertebrates. *Proc. Natl Acad. Sci. USA* **114**, E2375–E2384 (2017).
49. Brazeau, M. D. & Ahlberg, P. E. Tetrapod-like middle ear architecture in a Devonian fish. *Nature* **439**, 318–321 (2006).

# Methods

### Preparation and illustration of specimens

The specimens were collected from the Sosnovskiy Geological Monument, on the right bank of the river Izhma opposite Sosnogorsk Town (Komi Republic, Russia), during a series of field seasons from 2002 to 2012. The bulk of the material was collected during a large-scale excavation in 2009–2012, during which approximately 50 $m^2$ of the bone-bearing 'fish dolomite' bed was dug out and then broken into small blocks using hammers, chisels, an angle grinder, drill and portable jackhammer. Blocks that contained parts of the same bone were glued together. The bones were freed from the limestone matrix using dilute (7–10%) acetic acid, alternating with drying and covering with the consolidants PVB (before 2010) and paraloid B-72 (after 2010). The reconstructions of the skull and lower jaw were assembled by hand on the basis of photographs of individual bones, taken at appropriate angles. The reconstruction of the pectoral girdle was produced by sticking together the right anocleithrum, cleithrum, clavicle and interclavicle of one individual, making a three-dimensional virtual model of the assembly using photogrammetry (Agisoft PhotoScan), and importing this model into 3-matic (Materialise), in which it was duplicated, mirrored and assembled into a complete girdle. The drawings of the girdle in Fig. 3 were traced directly from lateral and anterior projections of the model.

### Phylogenetic analysis

The phylogenetic position of *Parmastega* was evaluated with maximum parsimony and Bayesian-inference analyses applied to a data matrix of 26 taxa and 113 characters (Supplementary Information), on the basis of a recently published matrix[50] with the addition of four characters (character numbers 7, 27, 28 and 29). Before all analyses, we explored the occurrence of possible 'taxonomic equivalents'[51] by subjecting the matrix to safe taxonomic reduction using the Claddis package[52] in the R environment for statistical computing and graphics (https://cran.r-project.org). No taxon was identified as being suitable for safe deletion.

For all parsimony analyses, we used PAUP* version 4.0a (build 164)[53] with the following search settings. The 'collapse branch' option was enforced for branches that could possibly attain a minimum length of zero. Tree searches used a heuristic option with a tree bisection–reconnection branch-swapping algorithm, and saving no more than a single tree of length greater than or equal to 1 step in each replicate, and using a maximum of 5,000 random step-wise taxon addition replicates while holding a single tree in memory at each step. Following this initial round of tree searches, an additional branch-swapping round was conducted on all trees saved in the memory—this time with the option of saving multiple trees in effect. This second round of tree searches was repeated ten times. No shorter or additional trees were found at the end of this second round in any of the parsimony analyses. Three analyses were carried out under maximum parsimony, each with the settings specified above.

In the first analysis, all characters were treated as unordered and of equal unit weight. We obtained 23 shortest trees at 278 steps, with an ensemble consistency index of 0.5 (0.4908 excluding 5 parsimony-uninformative characters), an ensemble retention index of 0.6911 and an ensemble rescaled consistency index of 0.3456. A permutation-tail probability test[54] using 1,000 replicates showed that the length of the shortest trees differed significantly from random ($P \sim 0.001$). The strict consensus tree (Extended Data Fig. 7a) was poorly resolved. The Adams consensus tree (Extended Data Fig. 7b) had greater resolution, and placed *Parmastega* and *Elginerpeton* as the joint (unresolved) sister groups to all other tetrapods. The agreement subtree (a pruned topology that included only those taxa for which all most-parsimonious trees agreed upon mutual relationships) included 19 out of the 26 original taxa (Extended Data Fig. 7e): *Acanthostega*, *Dendrerpeton*, *Densignathus*, *Elginerpeton*, *Greererpeton*, *Ossinodus* and *Tantallognathus* were deleted. The node support value was evaluated via bootstrapping[55] and jackknifing[56] in PAUP*, in each case using 50% character resampling and 50,000 random resampling replicates with the fast step-wise addition. In both cases, very few nodes received support—namely post-*Panderichthys* taxa, post-elpistostegalian taxa, baphetids and a clade of *Eoherpeton* plus *Proterogyrinus*.

In the second analysis, characters were re-weighted by the largest values of their rescaled consistency indexes from the initial analysis. PAUP* yielded a single tree (Extended Data Fig. 7c) that was 112.3561 steps long, with an ensemble consistency index of 0.6804 (0.6655 excluding uninformative characters), an ensemble retention index of 0.8297 and an ensemble rescaled consistency index of 0.5645. This tree was three steps longer than the trees from the unweighted analysis and did not represent a significantly better fit for the data, in terms of tree length, than the unweighted trees (based upon Templeton, Kishino–Hasegawa, and winning-sites tests in PAUP*). The weighted analysis confirmed the status of *Parmastega* as the most-basal tetrapod.

In the third analysis, we used implied weighting[57], experimenting with different integer values for Goloboff's constant of concavity $K$. We ran analyses with $1 \leq K \leq 10$ (for example, ref. [58]). For each $K$ value, we saved all trees that were generated at the end of the analysis. The separate tree files obtained from all $K$-weighted analyses were stored in PAUP* after filtering out duplicated tree topologies. This process resulted in five $K$-weighted trees, which were summarized with a strict consensus (Extended Data Fig. 7f), an agreement subtree (Extended Data Fig. 7g) and an Adams consensus (Extended Data Fig. 7h). The agreement subtree included 22 taxa: *Densignathus*, *Elginerpeton*, *Metaxygnathus* and *Ossinodus* were deleted.

For the Bayesian inference analysis, we used MrBayes v. 3.2.6 (ref. [59]), with the following settings: variable coding; gamma-distributed rate model; $10^7$ generations and four chains; and discarding the first 25% of sampled trees. The convergence diagnostic was evaluated through inspection of the potential scale reduction factor values[60] output by MrBayes. These values approached or were identical to 1, indicating successfully convergent runs (Supplementary Information). Credibility values for nodes in the Bayesian results (Extended Data Fig. 7d) were moderate-to-strong for most nodes.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

In total, 132 specimens comprising 183 skeletal elements were collected during the entire period of excavations (2002–2012). One hundred and six specimens (all of them figured in Supplementary Table 1) have been deposited in the collection of the Institute of Geology, Komi Science Centre (Ural Branch of the Russian Academy of Sciences, Syktyvkar, Russia) under collection number IG KSC 705/. One specimen has been deposited in the Ukhta Local Museum under collection number ULM 2599. The IG KSC and ULM specimens are available for examination. Other specimens have been reserved for sharing with other museums. The Life Science Identifier for *Parmastega* is urn:lsid:zoobank.org:act:76B5BB03-42FE-4F46-A284-F95E973CEE96.


50. Chen, D. et al. A partial lower jaw of a tetrapod from "Romer's Gap". *Earth Env. Sci. Trans. R. Soc. Edinb.* **108**, 55–65 (2018).
51. Wilkinson, M. Majority-rule reduced consensus trees and their use in bootstrapping. *Mol. Biol. Evol.* **13**, 437–444 (1996).
52. Lloyd, G. T. Estimating morphological diversity and tempo with discrete character-taxon matrices: implementation, challenges, progress, and future directions. *Biol. J. Linn. Soc.* **118**, 131–151 (2016).
53. Swofford, D. L. *PAUP* Phylogenetic Analysis Using Parsimony (*and Other Methods)* Version 4 (Sinauer, 2003).
54. Wilkinson, M., Peres-Neto, P. R., Foster, P. G. & Moncrieff, C. B. Type 1 error rates of the parsimony permutation tail probability test. *Syst. Biol.* **51**, 524–527 (2002).
55. Felsenstein, J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791 (1985).

56. Farris, J. S., Albert, V. A., Källersjö, M., Lipscomb, D. & Kluge, A. G. Parsimony jackknifing outperforms neighbor-joining. *Cladistics* **12**, 99–124 (1996).
57. Goloboff, P. A. Estimating character weights during tree search. *Cladistics* **9**, 83–91 (1993).
58. Congreve, C. R. & Lamsdell, J. C. Implied weighting and its utility in palaeontological datasets: a study using modelled phylogenetic matrices. *Palaeontology* **59**, 447–462 (2016).
59. Ronquist, F. & Huelsenbeck, J. P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574 (2003).
60. Gelman, A. & Rubin, D. B. Inference from iterative simulation using multiple sequences. *Stat. Sci*. **7**, 457–472 (1992).
61. Blom, H. Taxonomic revision of the Late Devonian tetrapod *Ichthyostega* from East Greenland. *Palaeontology* **48**, 111–134 (2005).
62. Ahlberg, P. E., Friedman, M. & Blom, H. New light on the earliest known tetrapod jaw. *J. Vertebr. Paleontol*. **25**, 720–724 (2005).
63. Robinson, J. The Evolution of the Early Tetrapod Middle Ear and Associated Structures. PhD thesis, Univ. College London (2006).

**Acknowledgements** We thank Y. Gatovsky, A. Zhuravlev and D. Ponomarev for their support of the project, and the dig crews of the 2009–2012 excavations for all their hard work. A. Ivanov identified the first tetrapod mandible from Sosnogorsk, in the Chernyshov Collection. P.A.B. acknowledges the support of National Geographic Society grant 9099-12 and UNDP/GEF project no. 00059042. E.L. acknowledges the support of Latvian Council of Science grant Z-6153-110. P.E.A. acknowledges the support of a Wallenberg Scholarship from the Knut and Alice Wallenberg Foundation.

**Author contributions** P.A.B. initiated and directed the excavation programme at Sosnogorsk, which produced the material for the study. E.L. and P.E.A. participated in excavations. P.A.B. carried out all preparation, consolidation and photography of specimens. P.E.A. made the reconstructions of the skull, lower jaw and shoulder girdle. M.R. performed the phylogenetic analyses. P.E.A. made Figs. 1–3 and Extended Data Figs. 2, 4–9. P.A.B. made Extended Data Figs. 1, 3, as well as Supplementary Tables 1, 2. P.A.B, J.A.C., E.L., M.R. and P.E.A. participated in the interpretation of the material and the writing of the paper.

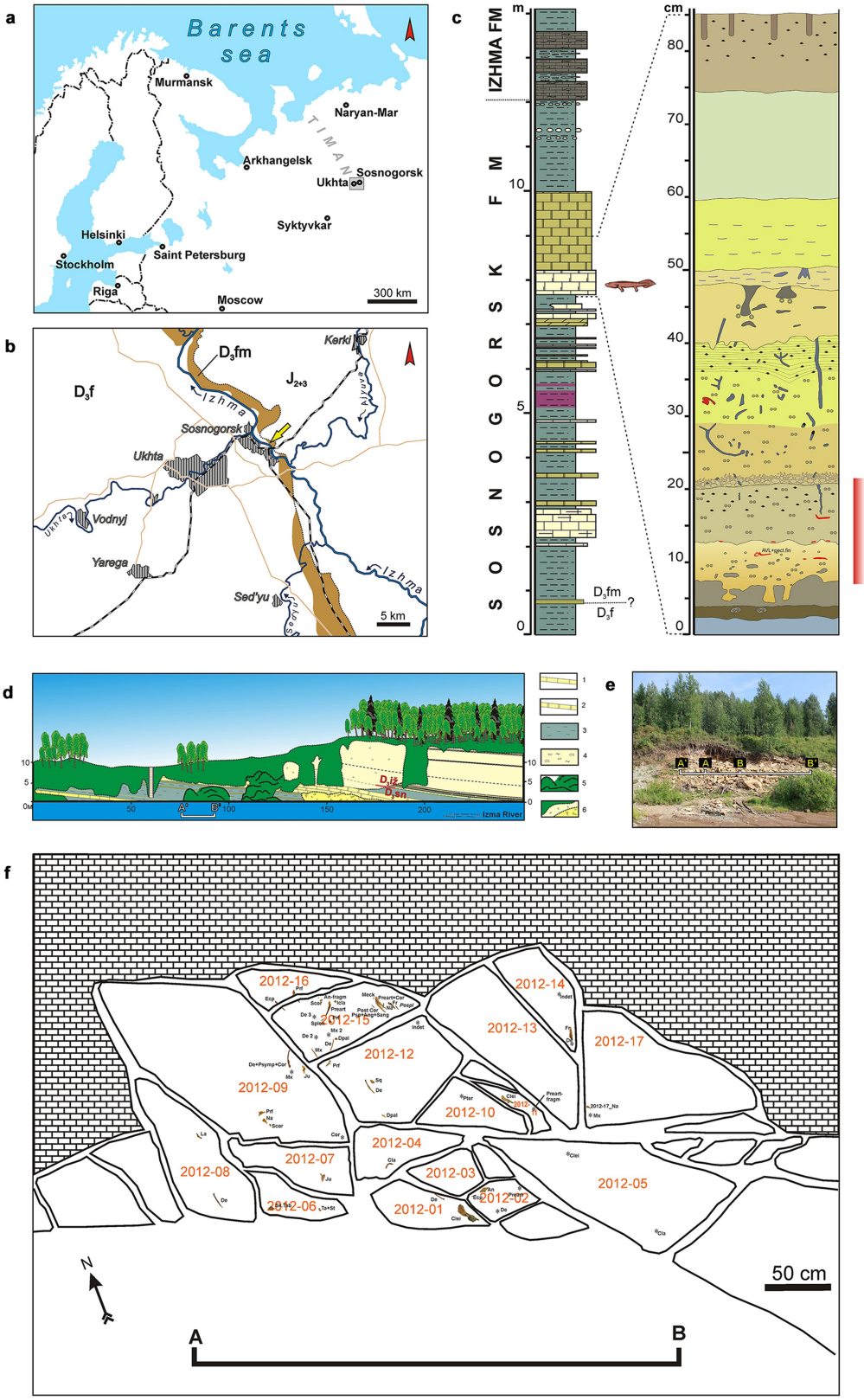
**Competing interests** The authors declare no competing interests.

**Additional information**

**Supplementary information** is available for this paper at https://doi.org/10.1038/s41586-019-1636-y.

**Correspondence and requests for materials** should be addressed to P.E.A.

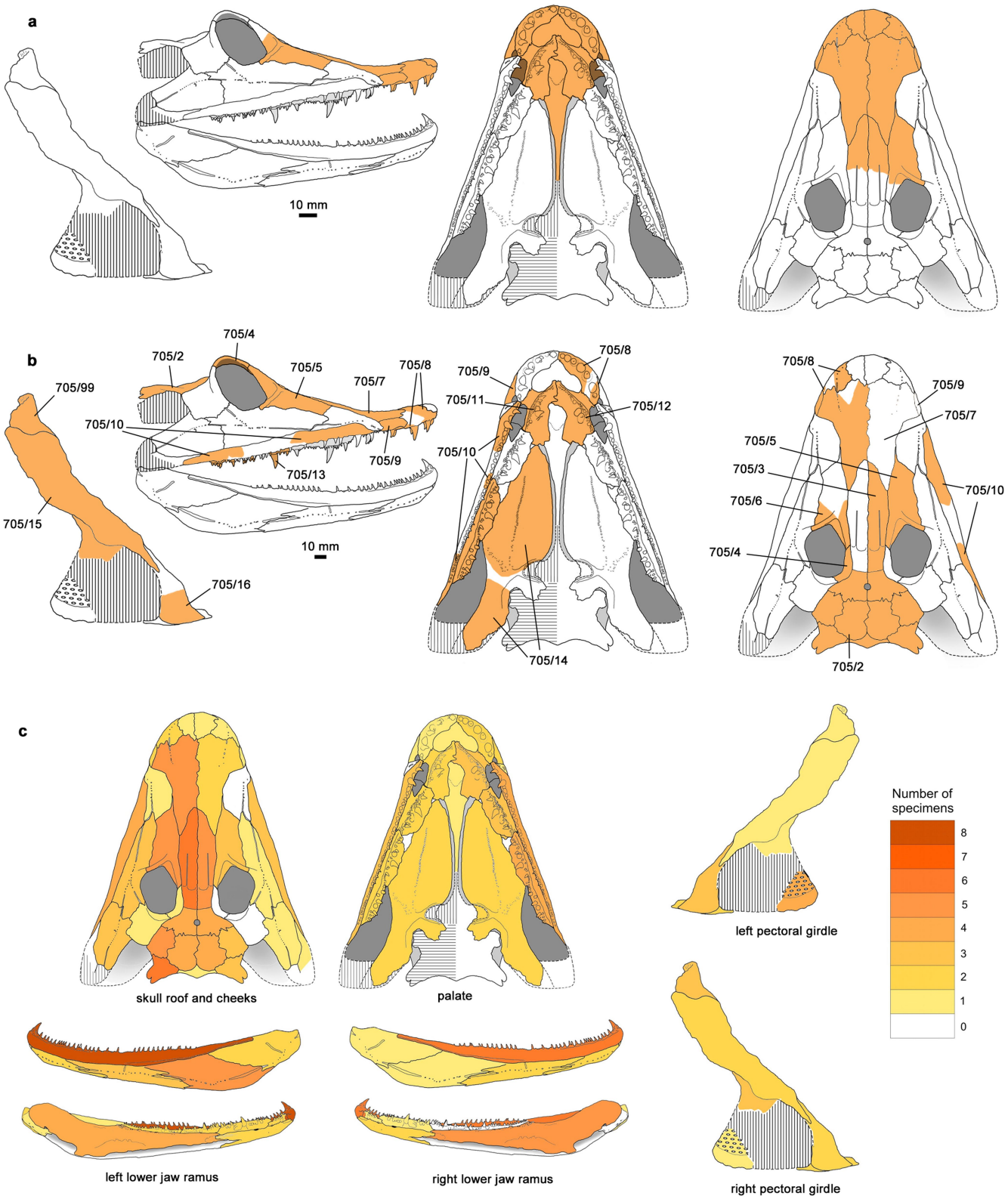
**Peer review information** *Nature* thanks Nadia Fröbisch and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



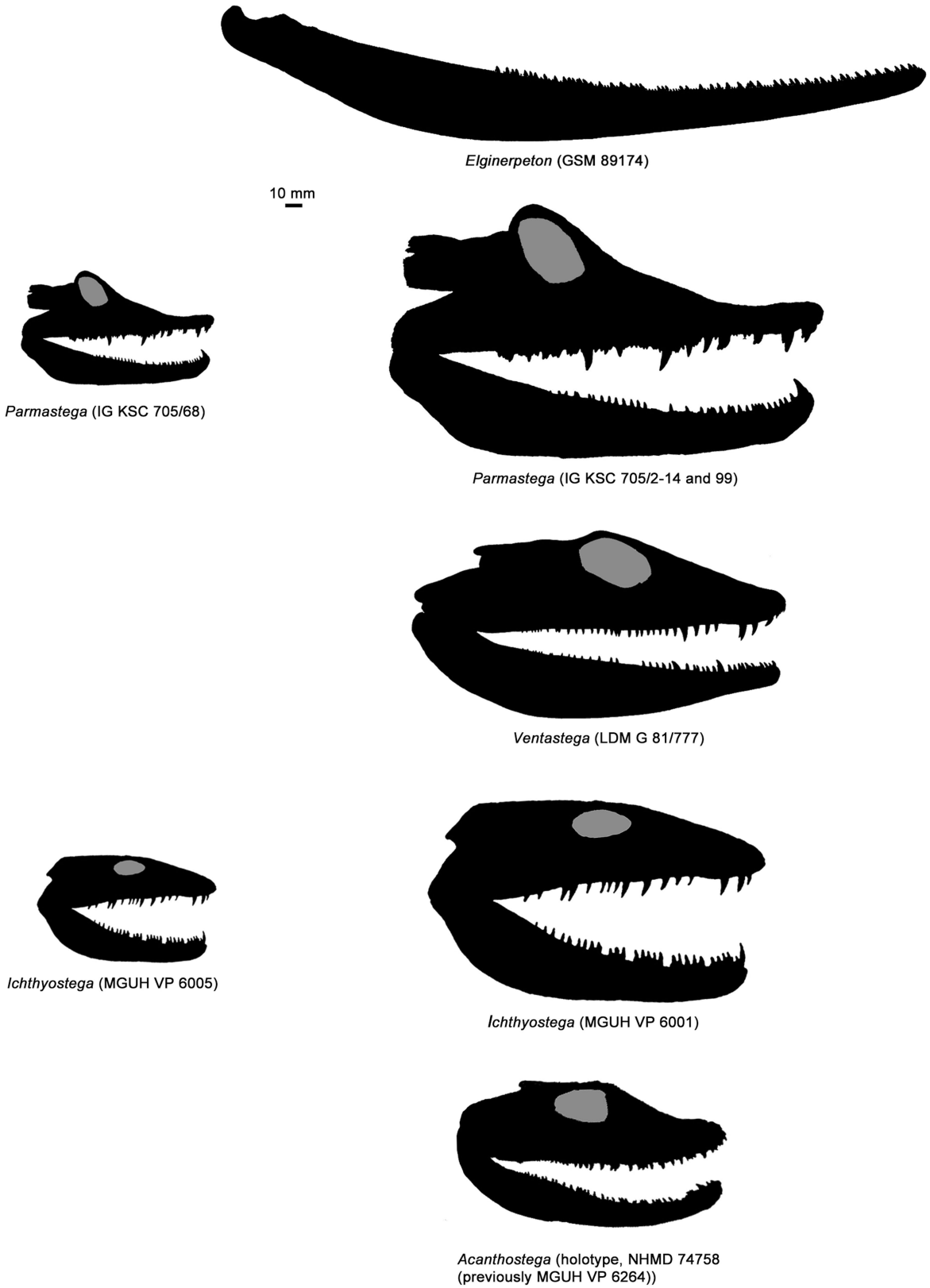

**Extended Data Fig. 1** | See next page for caption.

**Extended Data Fig. 1 | The distribution of *Parmastega* at the Sosnogorsk fossil site. a**, **b**, Maps of increasing resolution, showing the location of Sosnogorsk within northwest Russia. The box around Ukhta and Sosnogorsk in **a** indicates the region shown in **b**. In **b**, the brown belt that extends from north to south indicates the outcrop of Famennian ($D_3$fm) deposits in the region, and the yellow arrow points to the Sosnogorsk fossil site (Sosnovskiy Geological Monument). **c**, Stratigraphic column through the Sosnogorsk Formation, and part of the overlying marine Izhma Formation. Note the possible position of the Frasnian–Famennian boundary ($D_3$f–$D_3$fm) in the lower part of the Sosnogorsk Formation. The vertebrate-bearing part of the formation is shown in detail on the right; the tetrapod-bearing level is indicated with a red vertical bar. **d**, General view of outcrop no. 20 (Sosnovskiy Geological Monument) from the opposite bank of the Izhma River. 1, limestone; 2, dolomite; 3, clay; 4, nodular limestone; 5, scree; and 6, landslide. $D_3$sn, Sosnogorsk Formation, $D_3$iž, Izhma Formation. The distance A′–B′ indicates the area of the main excavation that took place in 2010–2012. **e**, Main excavation. The distance A–B indicates the area in which all of the tetrapod bones were found, during the excavation in 2012. The photograph was taken on 2 August 2012. **f**, Sketch map of the main excavation (2012), showing the distribution of tetrapod bones within the bed. The cluster numbers are indicated in orange. The background maps in **a** and **b** were taken from https://yandex.ru/maps; the geological features of **b** were taken from the open-access State Geological Map at https://vsegei.ru/.

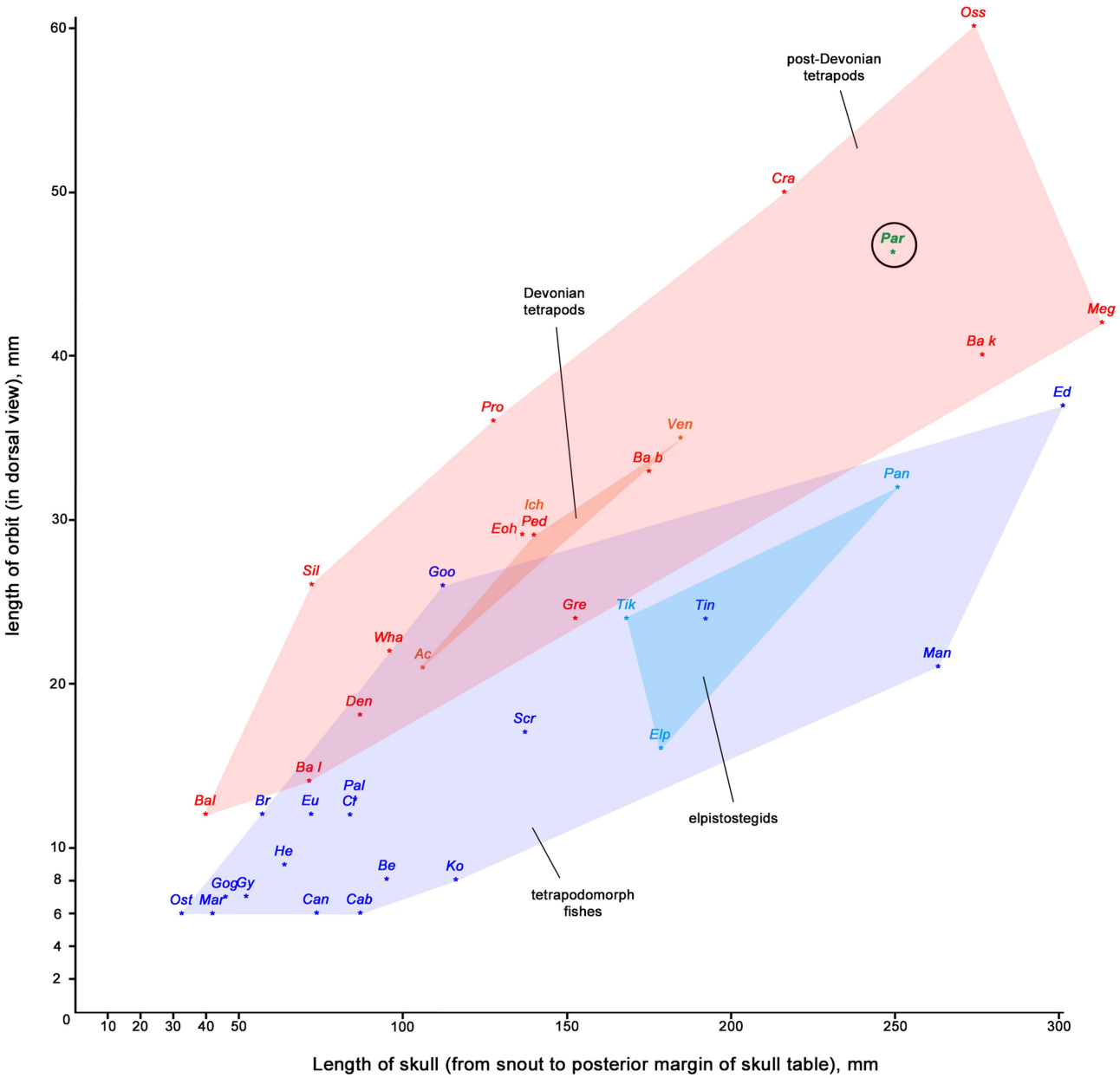

**Extended Data Fig. 2 | Frontal bones of *Parmastega*.** This figure shows all of the complete and near-complete frontals of *Parmastega* (eight out of nine known frontals) to scale, oriented with anterior at the top and aligned on the centre of radiation (horizontal line). The right frontals have been reversed so that all bones have the appearance of left frontals. From left to right, the specimens are IG KSC 705/3 (reversed), IG KSC 705/40, IG KSC 705/44 (reversed), IG KSC 705/43, IG KSC 705/45, IG KSC 705/18 (reversed), IG KSC 705/42 and IG KSC 705/41. Scale bar, 10 mm.



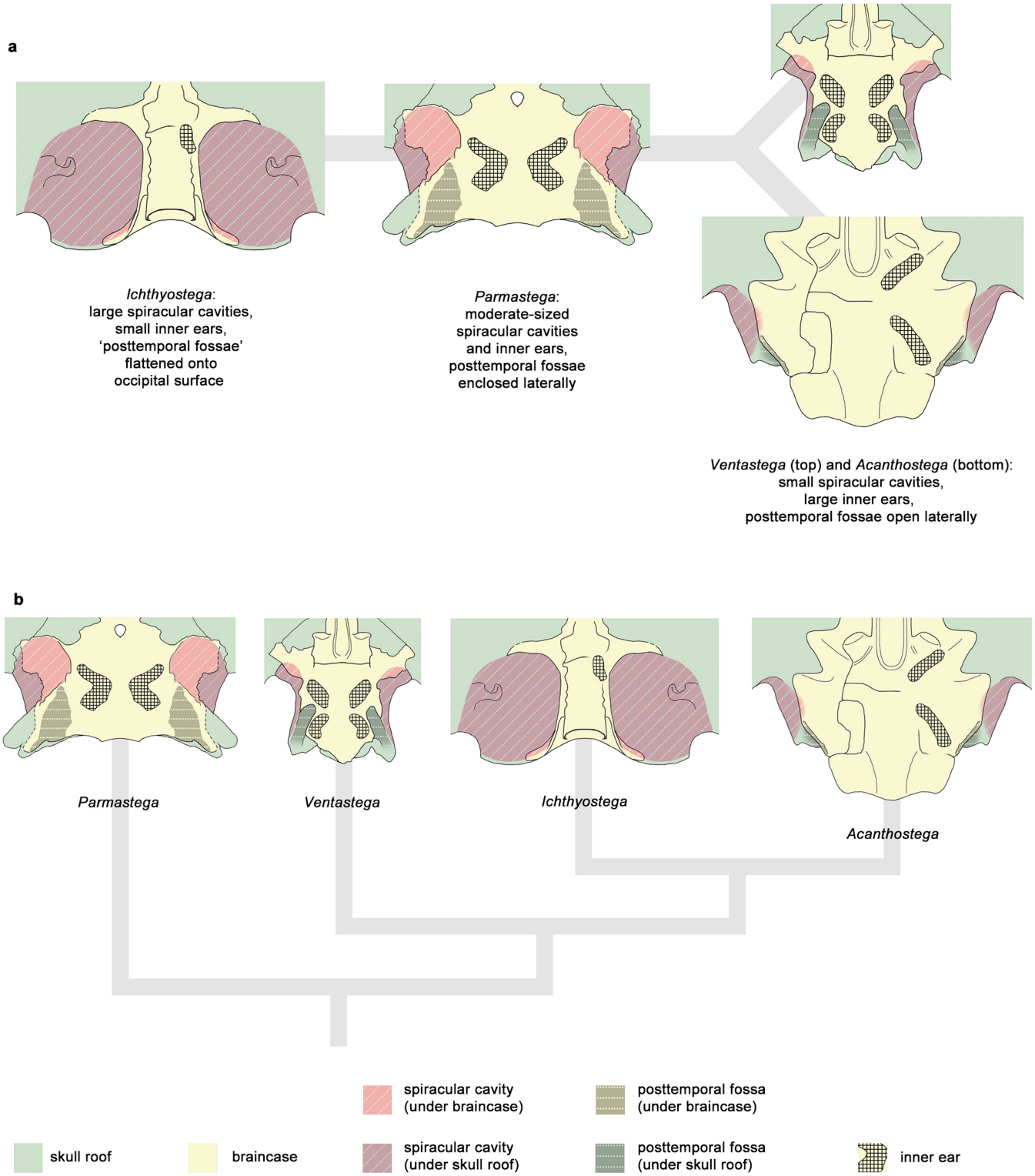

**Extended Data Fig. 3 | Bone associations. a**, **b**, Diagrammatic images showing the associated bones (in orange) of two individual skulls. **a**, The holotype IGKSC 705/1. **b**, The largest individual, IGKSC 705/2–705/14 and IGKSC 705/99. In the lateral view of **b**, the preserved frontal and nasal are shown (even though they are in fact on the other side of the skull). **c**, Diagrammatic representation of the number of specimens of different bones in the sample.



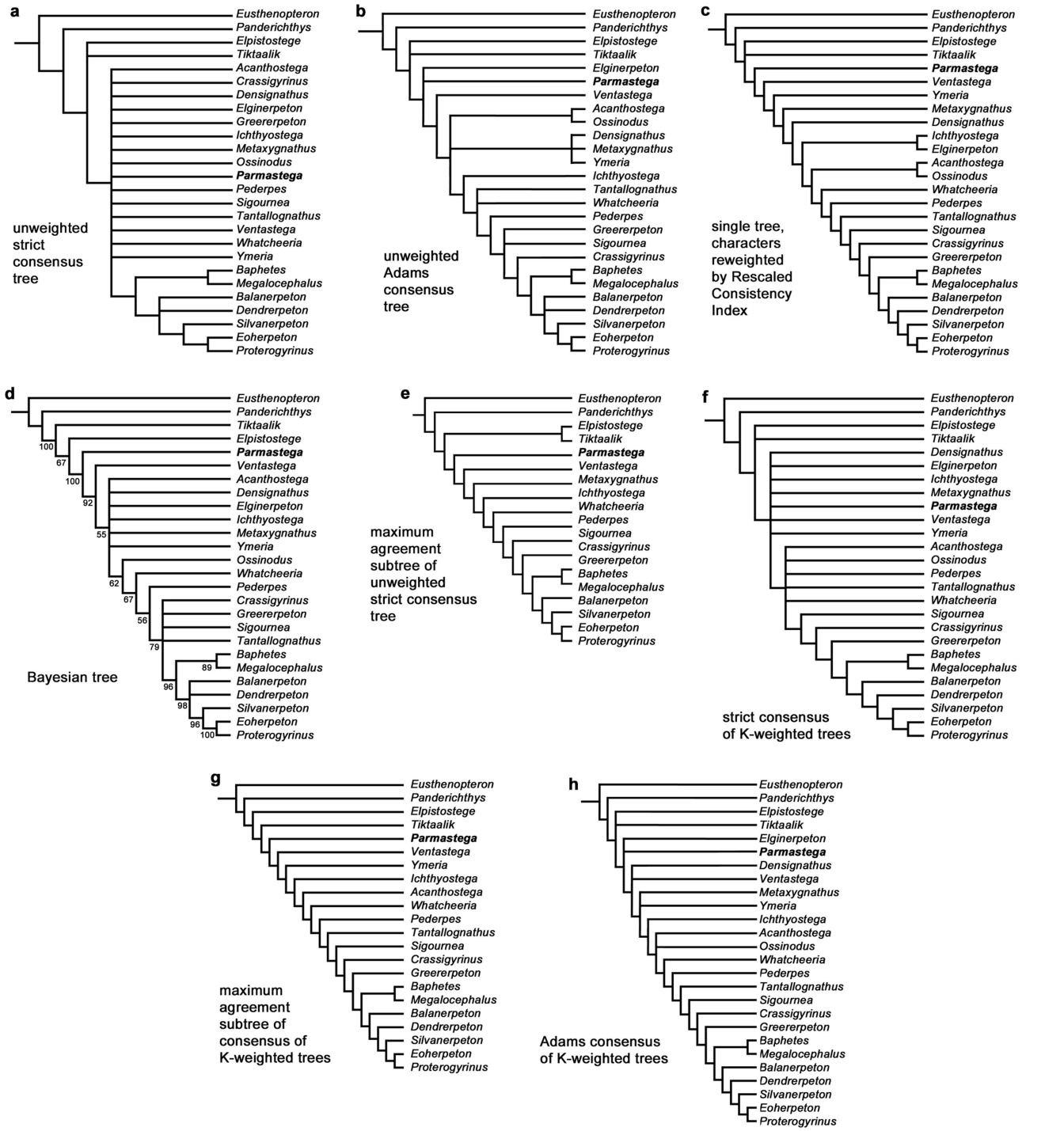

**Extended Data Fig. 4 | Size and shape of Devonian tetrapods.** Silhouette reconstructions of the heads of known, reconstructable Devonian tetrapods. Reconstructions are drawn to the same scale. The lower jaw of *Elginerpeton*—the largest known Devonian tetrapod, and for which the skull cannot be reconstructed—is also included. All reconstructions except for *Acanthostega* are assembled from more than one specimen; the specimen numbers indicate the specimen used to determine the scale. The right-hand column shows the largest known individuals. The left-hand column shows the smallest individuals of *Parmastega* (all from Sosnogorsk) and *Ichthyostega* (based on the entire East Greenland collection, reviewed in ref. [61]). Note the similarity in size range despite the very different nature of the samples. *Ventastega* and *Acanthostega* show narrow size ranges, which are not illustrated. Reconstructions modified from the following sources: *Ichthyostega*, ref. [19]; *Acanthostega*, ref. [31]; *Ventastega*, ref. [32]; *Elginerpeton*, ref. [62].



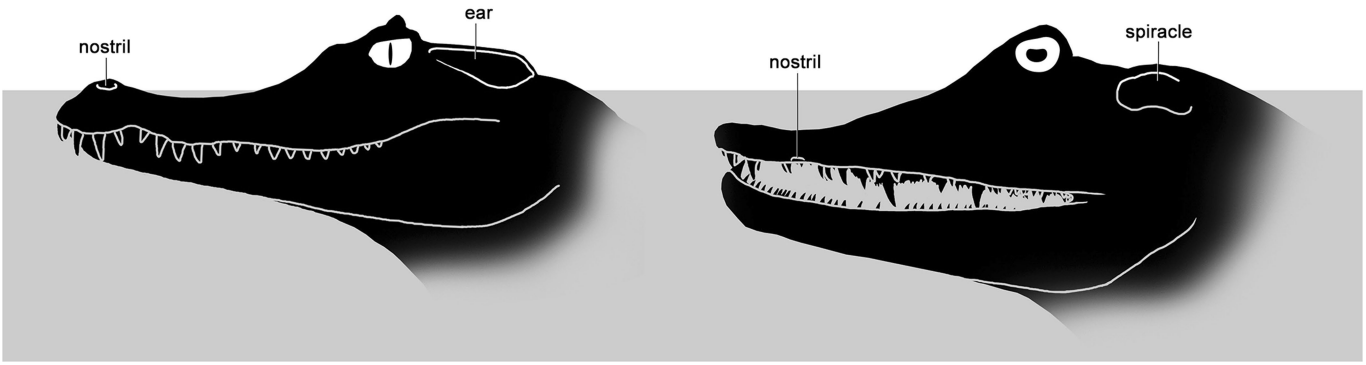

**Extended Data Fig. 5 | Relative orbit size.** Plot of orbit length versus skull length for a range of tetrapodomorph fishes, elpistostegids, Devonian tetrapods and post-Devonian tetrapods. Data are taken from ref. [47], except *Parmastega*, which is based on the largest known individual (Extended Data Fig. 3). Post-Devonian tetrapods from ref. [47] not included in our phylogenetic analysis are not shown. *Ac*, *Acanthostega*; *Ba b*, *Baphetes bohemicus*; *Ba k*, *Baphetes kirkbyi*; *Ba l*, *Baphetes lintonensis*; *Bal*, *Balanerpeton*; *Be*, *Beelarongia*; *Br*, *Bruehnopteron*; *Cab*, *Cabonnichthys*; *Can*, *Canowindra*; *Cl*, *Cladarosymblema*; *Cra*, *Crassigyrinus*; *Den*, *Dendrerpeton*; *Ed*, *Edenopteron*; *Elp*, *Elpistostege*; *Eoh*, *Eoherpeton*; *Eu*, *Eusthenopteron*; *Gog*, *Gogonasus*; *Goo*, *Gooloogongia*; *Gre*, *Greererpeton*; *Gy*, *Gyroptychius*; *He*, *Heddleichthys*; *Ich*, *Ichthyostega*; *Ko*, *Koharalepis*; *Man*, *Mandageria*; *Mar*, *Marsdenichthys*; *Meg*, *Megalocephalus*; *Oss*, *Ossinodus*; *Ost*, *Osteolepis*; *Pal*, *Palatinichthys*; *Pan*, *Panderichthys*; *Par*, *Parmastega*; *Ped*, *Pederpes*; *Pro*, *Proterogyrinus*; *Scr*, *Screbinodus*; *Sil*, *Silvanerpeton*; *Tik*, *Tiktaalik*; *Tin*, *Tinirau*; *Ven*, *Ventastega*; *Wha*, *Whatcheeria*.



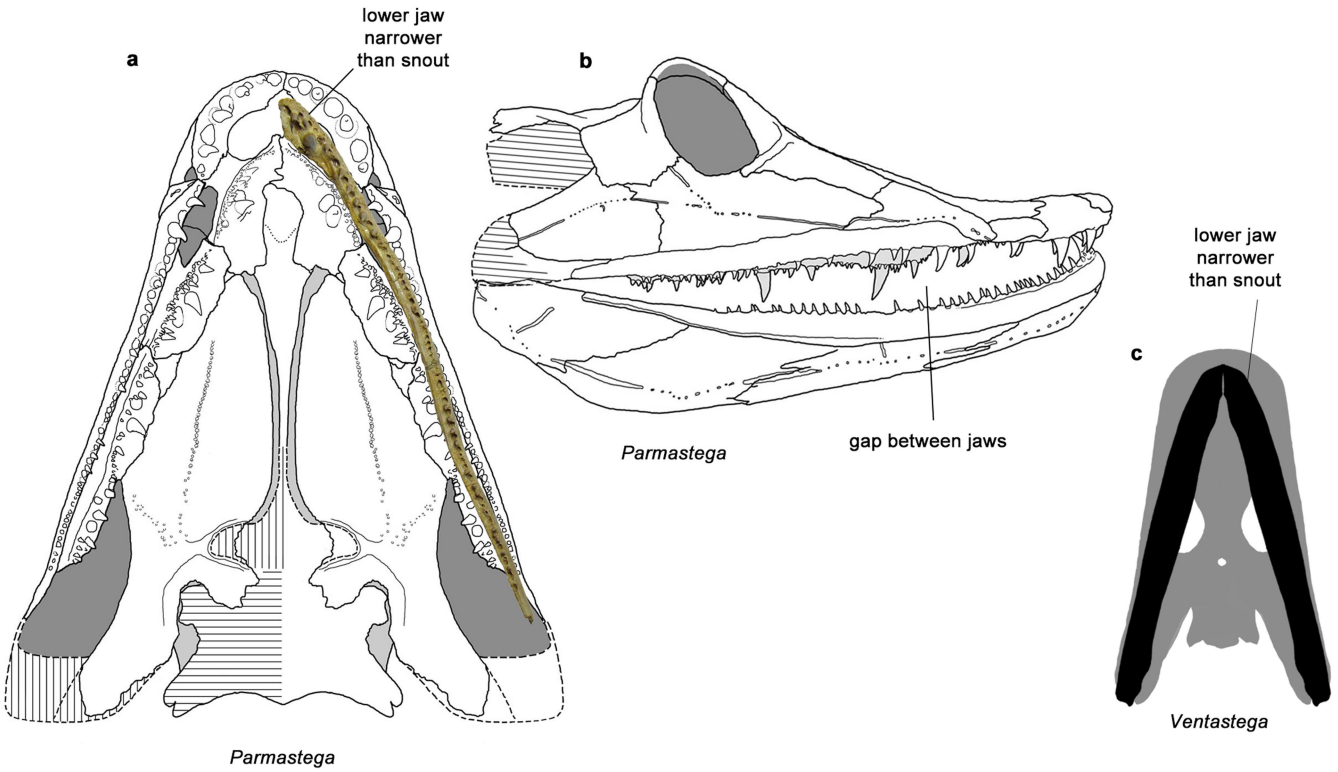

**Extended Data Fig. 6 | Otoccipital morphologies of Devonian tetrapods. a**, Comparative diagram of the otoccipial regions of *Parmastega*, *Ichthyostega* (new reconstruction, based on data from refs. [18,20]), *Ventastega* (modified from ref. [32]) and *Acanthostega* (modified from ref. [20], semicircular canals modified from ref. [63]) in ventral view. The basioccipital–exoccipital complex is preserved only in *Ichthyostega* and *Acanthostega*; in these genera the inner ear is shown only on one side. Drawings are scaled to the same length from pineal region to posterior margin of otic capsule. The inner ear is represented by the grooves for the anterior and posterior oblique semicircular canals, except in *Ichthyostega* in which it is represented by the sacculus (modified from ref. [20]). The braincases are arranged by morphological similarity, so that a minimum number of transformations are required along each branch. **b**, Consensus phylogeny from the analyses presented in this paper. The phylogenetic topology does not match the similarity dendrogram.




**Extended Data Fig. 7 | Phylogenetic analysis. a**, Unweighted strict-consensus tree. **b**, Unweighted Adams consensus tree. **c**, Single tree resulting from reweighting characters by the rescaled consistency index. **d**, Bayesian tree, with credibility values at nodes. **e**, Maximum-agreement subtree of unweighted parsimony analysis. **f**, Strict consensus of *K*-weighted trees. **g**, Maximum-agreement subtree of *K*-weighted parsimony analysis. **h**, Adams consensus of all trees from all *K*-weighted analyses.




**Extended Data Fig. 8 | *Parmastega* and caiman.** Comparison in left lateral view of spectacled caiman (*Caiman crocodilus*) on the left and *Parmastega* on the right, drawn to the same size, showing the inferred similar cruising posture at the surface. Note the difference in the positions of the nostrils. The caiman image is based on a computed tomography scan of a skull in the Digimorph Archive (http://www.digimorph.org/specimens/Caiman_crocodilus/).




**Extended Data Fig. 9 | Fit of dentary against upper jaw. a**, Dentary of *Parmastega* (IG KSC 705-67) fitted against palatal reconstruction to show the difference in curvature between the spade-shaped snout and the relatively straight dentary. **b**, Lateral view of skull reconstruction of *Parmastega* with closed mouth, showing mismatch in curvature between upper and lower jaws. **c**, Composite reconstruction of *Ventastega*, superimposing lower jaw rami (from ref. [30]) on skull reconstruction (from ref. [32]), showing shape relationship similar to **a**. Not to scale.

nature research

Corresponding author(s): Per Erik Ahlberg

Last updated by author(s): Jun 26, 2019

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| n/a | Confirmed |
|---|---|
| ☒ | ☐ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☒ | ☐ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☒ | ☐ The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☒ | ☐ A description of all covariates tested |
| ☒ | ☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☒ | ☐ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☒ | ☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☐ | ☒ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒ | ☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |

*Our web collection on statistics for biologists contains articles on many of the points above.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | Materialise 3-matic Research 12.0 (software for manipulating three-dimensional virtual objects in space; used for constructing model of pectoral girdle of Parmastega). |
| Data analysis | PAUP* version 4.0a and MrBayes version 3.2.6 (for phylogenetic analysis) |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All specimens figured and described in the paper are accessioned to the Institute of Geology, Komi Science Centre, Ural Branch of the Russian Academy of Sciences, Syktyvkar, Russia, and are deposited there. The accession code is IG KSC 705/. All the specimens are available for examination.

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

# Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Study description | Description of fossil material of the Devonian tetrapod Parmastega aelidae. |
| Research sample | All known specimens of this taxon. |
| Sampling strategy | We excavated the fossil locality (Sosnovskiy Geological Monument, at Sosnogorsk on the bank of the Izhma River) and collected all the fossils we could find. The fossils were freed from the rock with dilute acetic acid by Pavel Beznosov. |
| Data collection | The primary interpretation of the fossils and the assembly of the reconstruction were undertaken by Pavel Beznosov and Per Ahlberg, working with the specimens in Syktyvkar during a series of visits by Per Ahlberg. |
| Timing and spatial scale | Excavations were carried out during 2002-2012. |
| Data exclusions | No data were excluded. |
| Reproducibility | Not applicable. |
| Randomization | Not applicable. |
| Blinding | Not applicable. |

Did the study involve field work? ☒ Yes ☐ No

## Field work, collection and transport

| | |
|---|---|
| Field conditions | Typical summer weather in northern Russia. The weather conditions had no impact on data gathering. |
| Location | Sosnovskiy Geological Monument, Sosnogorsk, right bank of Izhma River, Komi Republic, Russia. |
| Access and import/export | The fieldwork was carried out by the Geological Institute of the Komi Science Center, Uralian Branch of the Russian Academy of Sciences, in accordance with local and national regulations. The material was not exported from Russia. |
| Disturbance | Blocks of limestone were removed from the riverbank. The annual ice-melt and spring flood of the Izhma River vigorously scours the banks and soon removes any trace of human disturbance. |

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials & experimental systems

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ Antibodies |
| ☒ | ☐ Eukaryotic cell lines |
| ☐ | ☒ Palaeontology |
| ☒ | ☐ Animals and other organisms |
| ☒ | ☐ Human research participants |
| ☒ | ☐ Clinical data |

## Methods

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ ChIP-seq |
| ☒ | ☐ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |

## Palaeontology

| | |
|---|---|
| Specimen provenance | The specimens come from the Sosnovskiy Geological Monument, Sosnogorsk, right bank of Izhma River, Komi Republic, Russia, and were collected by the Geological Institute of the Komi Science Center, Uralian Branch of the Russian Academy of Sciences in accordance with local and national regulations. |
| Specimen deposition | Geological Institute of the Komi Science Center, Uralian Branch of the Russian Academy of Sciences |
| Dating methods | Not applicable |

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Article

# The landscape of somatic mutation in normal colorectal epithelial cells

https://doi.org/10.1038/s41586-019-1672-7

Received: 11 September 2018

Accepted: 11 September 2019

Published online: 23 October 2019

Henry Lee-Six[1], Sigurgeir Olafsson[1], Peter Ellis[1], Robert J. Osborne[1], Mathijs A. Sanders[1,2], Luiza Moore[1], Nikitas Georgakopoulos[3,4], Franco Torrente[5], Ayesha Noorani[6,7], Martin Goddard[8], Philip Robinson[1], Tim H. H. Coorens[1], Laura O'Neill[1], Christopher Alder[1], Jingwei Wang[1], Rebecca C. Fitzgerald[6,7], Matthias Zilbauer[5,9], Nicholas Coleman[7,10], Kourosh Saeb-Parsy[3,4], Inigo Martincorena[1], Peter J. Campbell[1] & Michael R. Stratton[1]*

The colorectal adenoma–carcinoma sequence has provided a paradigmatic framework for understanding the successive somatic genetic changes and consequent clonal expansions that lead to cancer[1]. However, our understanding of the earliest phases of colorectal neoplastic changes—which may occur in morphologically normal tissue—is comparatively limited, as for most cancer types. Here we use whole-genome sequencing to analyse hundreds of normal crypts from 42 individuals. Signatures of multiple mutational processes were revealed; some of these were ubiquitous and continuous, whereas others were only found in some individuals, in some crypts or during certain periods of life. Probable driver mutations were present in around 1% of normal colorectal crypts in middle-aged individuals, indicating that adenomas and carcinomas are rare outcomes of a pervasive process of neoplastic change across morphologically normal colorectal epithelium. Colorectal cancers exhibit substantially increased mutational burdens relative to normal cells. Sequencing normal colorectal cells provides quantitative insights into the genomic and clonal evolution of cancer.

Sequencing of the genomes of over 20,000 cancers of several types has identified the repertoire of driver mutations in cancer genes that convert normal cells into cancer cells and revealed the mutational signatures of the underlying biological processes that generate somatic mutations[2,3]. Cancers are, however, end stages of an evolutionary process that operates within populations of cells, and commonly arise through the accumulation of several driver mutations that engender a series of clonal expansions. Understanding this progression has depended on the identification of somatic mutations in morphologically abnormal neoplastic proliferations that represent intermediate stages between normal cells and cancer cells[1].

As for most cancer types, the earliest stages of progression to colorectal cancer remain less well understood. The driver mutation that first sets a colorectal epithelial cell on the path to cancer is probably caused by mutational processes that also operate in normal cells and that we only understand to a limited extent. The nature and numbers of the earliest neoplastic clones with driver mutations—which conceivably are morphologically indistinguishable from normal cells—are similarly unclear. In large part, these deficiencies are due to the technical challenge of identifying somatic mutations in normal tissues, which are composed of myriad microscopic cell clones. Several different approaches have been adopted to address this challenge[4–14], and have revealed the signatures of common somatic mutational processes in normal cells of the small and large intestine, liver, blood, skin and nervous system. Thus far, however, studies have not been of sufficient scale to characterize variation in signature activity or detect processes that occur less frequently[4–14]. High proportions of normal skin, oesophageal and endometrial epithelial cells have been shown to be members of clones that already carry driver mutations[10,11,15,16], and large mutant clones have been detected in the blood[17–20]. The extent of this phenomenon in the colon, an organ with a high incidence of cancer, has not been investigated.

Colonic epithelium is a contiguous cell sheet organized into around 15 million crypts, each of which is composed of about 2,000 cells[21]. Towards the base of each crypt, a small number of stem cells are found, which are ancestral to the maturing and differentiated cells in the crypt[22]. These stem cells stochastically replace one another through a process of neutral drift[23,24], such that all stem cells—and thus all cells—in a crypt derive from a single ancestral stem cell that existed in recent years[25–27]. The somatic mutations that were present in this ancestor are thus found in all of the approximately 2,000 descendant cells and can be revealed by DNA sequencing of an individual crypt. These stem cells are thought to be the cells of origin of colorectal cancers[28]. To characterize the earliest stages of colorectal carcinogenesis, we examined somatic mutational burdens, mutational signatures, clonal dynamics and the frequency of driver mutations in normal colorectal epithelium by sequencing individual colorectal crypts.

[1]Wellcome Sanger Institute, Hinxton, UK. [2]Department of Hematology, Erasmus University Medical Center, Rotterdam, The Netherlands. [3]Department of Surgery, University of Cambridge, Cambridge, UK. [4]Cambridge NIHR Biomedical Research Centre, Cambridge Biomedical Campus, Cambridge, UK. [5]Department of Paediatric Gastroenterology, Hepatology and Nutrition, Addenbrooke's Hospital, Cambridge, UK. [6]Medical Research Council Cancer Unit, Hutchison/Medical Research Council Research Centre, University of Cambridge, Cambridge, UK. [7]Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK. [8]Department of Pathology, Papworth Hospital NHS Trust, Cambridge, UK. [9]Department of Paediatrics, University of Cambridge, Cambridge, UK. [10]Department of Pathology, University of Cambridge, Cambridge, UK. *e-mail: mrs@sanger.ac.uk

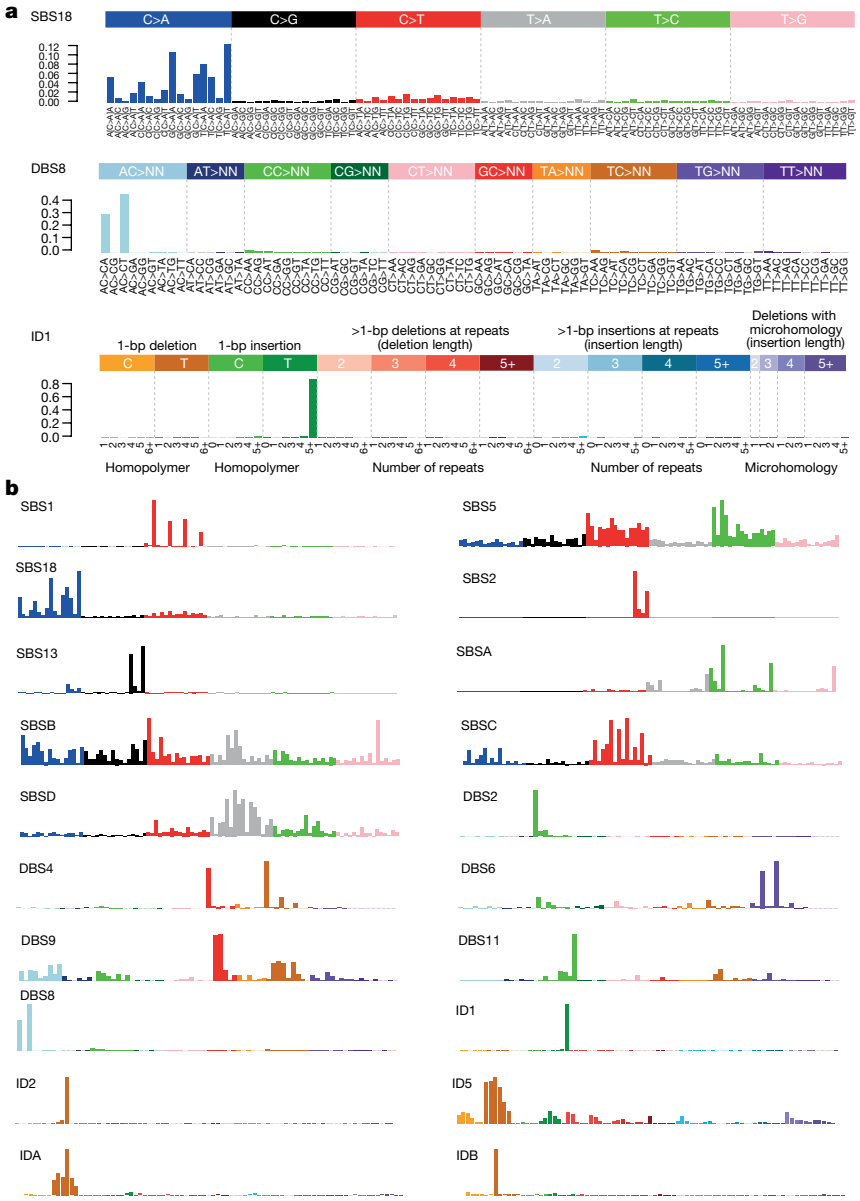



**Fig. 1 | Mutational signatures that are present in normal colon. a**, Examples of an SBS, a DBS and an ID signature (SBS18, DBS8 and ID1), showing the categories into which mutations are divided. Later figures are shown in the same format. **b**, The complement of signatures in normal colonic epithelium. Known signatures are labelled according to their nomenclature in the Pan Cancer Analysis of Whole Genomes (PCAWG)[2] and novel signatures are labelled with letters.

## Somatic mutations and mutational signatures

We used laser-capture microdissection to isolate 2,035 individual colonic crypts from the normal epithelium of 42 individuals aged 11 to 78, of whom 15 had a history of colorectal cancer and 27 did not (Methods, Supplementary Table 1), and sequenced their genomes. The distribution of variant allele fractions (VAFs) from whole-genome sequencing of 571 individual crypts showed that crypts were derived from a single ancestral stem cell (Extended Data Fig. 1d), and simulations indicated that about 90% of the mutations called were fully clonal (Supplementary Information). There was substantial variation in mutational burdens between individual crypts—for example, the mutational burden ranged from 1,508 to 15,329 for individuals in their sixties—and this was not obviously attributable to technical factors. To explore the biological basis of this variation we extracted mutational signatures and estimated the contribution of each to the mutational burden of every crypt (Methods, Supplementary Information).

Nine single-base substitution (SBS), six doublet-base substitution (DBS) and five small insertion and deletion (indel) (ID) mutational signatures were found. Of these, 14 closely matched (Methods) a known reference signature (SBS1, SBS2, SBS5, SBS13, SBS18, DBS2, DBS4, DBS6, DBS8, DBS9, DBS11, ID1, ID2 and ID5; nomenclature as described previously[2]) and six did not (SBSA, SBSB, SBSC, SBSD, IDA and IDB) (Fig. 1, Extended Data Figs. 2–4). Thus, new mutational signatures were extracted despite extensive previous analysis of cancers—perhaps owing to masking by the comparative complexity of the mixes of signatures that are present in cancer genomes.

## Ubiquitous mutational signatures

Eleven signatures (SBS1, SBS5, SBS18, DBS2, DBS4, DBS6, DBS9, DBS11, ID1, ID2 and ID5) were found in over 85% of crypts and are here termed 'ubiquitous'. All have been previously described[2].

SBS1 is characterized by C>T substitutions at NCG trinucleotides (the mutated base is underlined) and is probably a result of the deamination of 5-methylcytosine. Its mutational load correlated linearly with age (Fig. 2). There was, however, variation in SBS1 mutational burdens

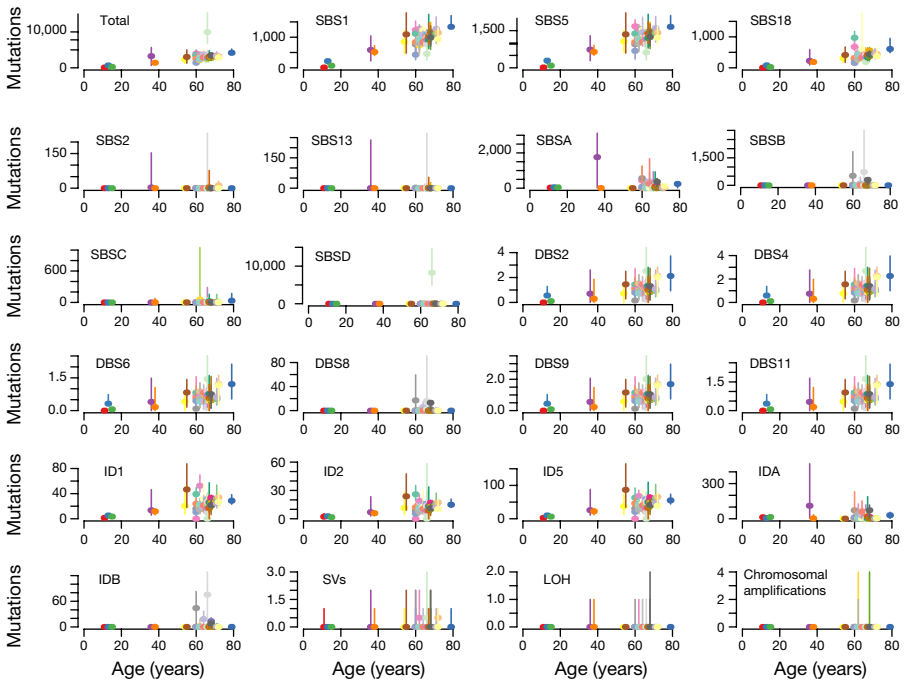



**Fig. 2 | Mutational burden versus age for every signature.** For every signature, the median (horizontal bar) and range (vertical bar) in mutational burden for all the crypts from each individual are shown. Each individual is coloured differently. $n = 445$ crypts from 42 individuals.

between crypts from the same individual ($P = 2.25 \times 10^{-27}$). This was, in part, due to different SBS1 mutation rates in different sectors of the colon, with mean rates across individuals of 16.8 mutations per year (95% confidence interval, 15.2–18.3) in the right colon (ascending and caecum), 16.1 (95% confidence interval, 14.4–17.5) in the transverse colon and 12.8 (95% confidence interval, 11.1–14.4) in the left colon (descending and sigmoid). The SBS1 mutation rate in the terminal ileum was 12.7 (95% confidence interval, 10.6–14.9) (Supplementary Information). SBS5 is a relatively featureless signature of unknown cause and SBS18 is characterized by C>A mutations, which may be caused by damage to DNA by reactive oxygen species[29,30]. The mutational burdens of these two signatures correlated with age, with the same ordering of sector differences as SBS1 ($P = 9.89 \times 10^{-26}$ for SBS5; $P = 5.43 \times 10^{-22}$ for SBS18). Even after taking anatomical location and age into account, differences in mutational burden remained between different crypts, notably for SBS18 (Fig. 2, Extended Data Fig. 9, Extended Data Fig. 6al). Combining ubiquitous SBS mutational signatures, and averaging over anatomical sites, the rate of mutation was 43.6 mutations per year, which is comparable with previous estimates[4].

DBS2, DBS4, DBS6, DBS9 and DBS11 were tightly correlated in all colonic crypts. ID1, ID2 and ID5—which are characterized by indels of a single T and may be the consequence of slippage during DNA replication—all accumulated linearly with age, with the same order of sector differences as SBS1 ($P = 1.66 \times 10^{-5}$ for ID1, $P = 4.53 \times 10^{-6}$ for ID2 and $P = 4.53 \times 10^{-6}$ for ID5) (Supplementary Information, Extended Data Fig. 5).

The correlations of ubiquitous signatures with age indicate that the mutational processes that underlie them operate throughout life, in all individuals and all colorectal stem cells. However, the results also suggest that differences in physiology and/or microenvironment (and, potentially, the age of the most recent common ancestor of the crypts[27]) between different sectors of the colon cause measurable differences in somatic mutation rates.

## Sporadic mutational signatures

Nine signatures (SBS2, SBS13, SBSA, SBSB, SBSC, SBSD, DBS8, IDA and IDB) were present only in a subset of individuals and/or a subset of crypts and are here termed 'sporadic'. All were novel, except for SBS2, SBS13 and DBS8. SBS2 and SBS13 are characterized by C>T and C>G mutations at TCN, are probably caused by cytidine deaminases of the APOBEC family and usually occur together[31,32]. They were present in only two crypts (a colonic crypt (Extended Data Fig. 6ai) and an ileal crypt (Extended Data Fig. 6ao) from different individuals), occurring together and each accounting for over 150 mutations. To our knowledge, this is the first report that DNA editing of the human genome by APOBEC cytidine deaminases occurs in normal cells in vivo. The sequence context of these mutations in normal colon suggests that APOBEC3A is the major contributing enzyme[33].

Four SBS signatures that do not match the reference set, SBSA, SBSB, SBSC and SBSD, were found in normal colorectal cells (SBSA has recently been reported in an oral squamous carcinoma[34]). SBSA is characterized by T>C mutations at ATA, ATT and TTT, and T>G mutations at TTT. Its mutational burden correlated closely with that of IDA, in which single T deletions in short runs of T bases (with a modal average of four) predominate—suggesting that these two signatures are a result of the same underlying mutational process. SBSA was detectable in 29 out of 42 individuals and often accounted for thousands of mutations in just a subset of crypts. It clustered spatially in the colon, with crypts from the same biopsy exhibiting the signature even though the mutations themselves were not shared (Supplementary Information, Extended Data Fig. 9). Around 2.5-fold more T>C mutations occurred when the T base was on the transcribed than on the untranscribed strand. Transcriptional-strand bias is often caused by transcription-coupled nucleotide excision repair acting on DNA that has been damaged by exogenous exposures that cause covalently bound bulky adducts; however, it can also occur as a result of transcription-coupled DNA damage[35]. Assuming that one of these two possibilities is the case, damage to adenine underlies SBSA. To investigate the timing of SBSA, we constructed phylogenetic trees of mutations and established the mutational signatures in each branch (Fig. 3, Extended Data Fig. 6). SBSA was confined to early branches of these phylogenies (when these were available for analysis) (Fig. 3b, Extended Data Fig. 6f, h, z, aa, am, ao, aq). Using the number of SBS1 mutations as indicators of real time, the mutational process that underlies SBSA appears to be active before an individual reaches

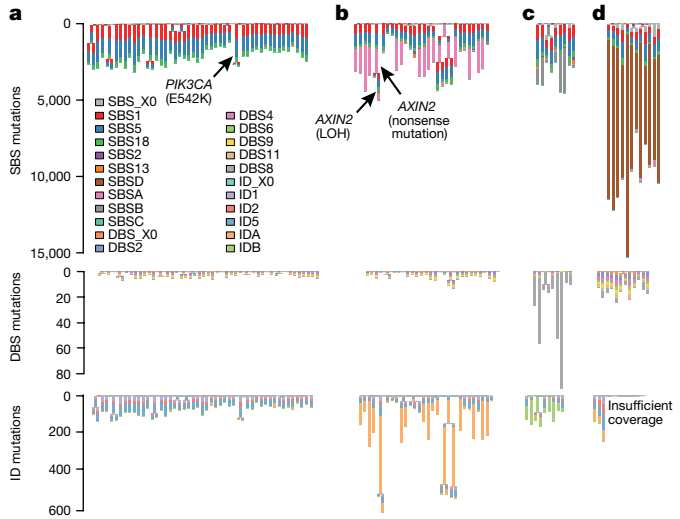



**Fig. 3 | Crypt phylogenies. a–d**, For selected individuals (**a**–**d**), each phylogeny is shown three times; branch lengths are proportional to mutation counts for SBS (top), DBS (middle) and ID (bottom) signatures. A stacked bar plot of the signatures that contribute to each branch is superimposed. The ordering of signatures along branches is for visualization purposes. 'X0' indicates mutations that could not confidently be assigned to any signature. Phylogenies for all individuals are shown in Extended Data Fig. 6. Selected phylogenies are dominated by ubiquitous signatures (**a**), SBSA and IDA (**b**), SBSB, DBS8 and IDB (**c**) or SBSD for the individual treated with chemotherapy (**d**).

10 years of age (Supplementary Methods, Extended Data Fig. 6aq). SBSA may therefore be caused by an extrinsic, locally acting and patchily distributed mutagenic insult that occurs during childhood.

SBSB was characterized by C>T mutations at ACA, T>A at CTN and T>G at GTG, and was present in subsets of crypts from four individuals (Extended Data Fig. 6e, aa, ai, aj). It accounted for variable numbers of substitutions, with a maximum of 3,002 in one crypt (Fig. 3c, Extended Data Fig. 6ai). In the two individuals in whom SBSB could be timed (Extended Data Fig. 6aa, ai, aq), it appeared—as with SBSA—to be most active in the first decade of life. SBSB correlated with DBS8 and IDB (Fig. 3c, Extended Data Fig. 9), suggesting that they are caused by the same underlying mutational process. DBS8 is composed of AC>CA and AC>CT mutations and has previously been reported in rare hypermutated cancers with no obvious cause[2]. IDB is dominated by deletion of a single T that has no other T bases surrounding it.

SBSC is characterized by one C>T mutation in CC dinucleotides. It primarily affected three crypts (with 1,050, 827 and 695 mutations, respectively) from the left colon of one individual with an unremarkable history (Extended Data Fig. 9, Extended Data Fig. 6m, Supplementary Table 1).

All crypts from a 66-year-old man carried many thousands of mutations of SBSD (Fig. 3d, Extended Data Fig. 6ap), which is characterized by T>A substitutions with a transcriptional-strand bias that is compatible with damage to adenine. This individual had been treated with multiple chemotherapeutic agents (cyclophosphamide, doxorubicin, vincristine, prednisolone, chlorambucil, bleomycin and etoposide) for lymphoma and subsequently developed caecal adenocarcinoma. SBSD resembles SBS25 (cosine similarity, 0.9), which has previously been found in Hodgkin's lymphoma cell lines from two patients who were treated with chemotherapy[31,36]. To our knowledge, this is the first time that the mutational consequences of chemotherapy have been demonstrated in normal human cells in vivo. The mutational burden in the colorectal epithelium of this individual was three- to fivefold higher than that expected for his age—thus by extrapolation equivalent to that of a 200–300-year-old.

## Copy-number changes and structural variants

Copy-number changes and/or structural variants were found in 80 out of 449 (18%) crypts that could be evaluated. Five crypts exhibited eight whole-chromosome copy-number increases that affected the same three chromosomes—3, 7 and 9—as well as the X chromosome (Extended Data Fig. 7a). Thus, copy-number increases clustered in certain crypts and tended to affect certain chromosomes. No whole-chromosome losses were observed. Arm-level copy-number increases that affect chromosome 7 are common in colorectal cancers[37] and adenomas[38]. Copy-number increases in chromosomes 3 and 9 are seen in colorectal cancers, but are almost as frequently deleted[37]. Copy-neutral loss of heterozygosity (CN-LOH) was observed in 12 crypts, and affected chromosomes 1p, 6p, 7p, 8q, 9q, 10q (twice), 17p, 17q, 18q, 21q and 22q (Extended Data Fig. 7c). CN-LOH is frequently observed in colorectal cancers, although the specific changes that we observe here are not recurrent features[39]. Five copy-number changes could be timed and all were estimated to have occurred in adulthood (Extended Data Fig. 7b). Two changes that affected the same crypt appeared to be synchronous (Supplementary Information). An analysis of structural variants detected 48 large deletions, 18 tandem duplications, 4 translocations and 2 inversions (Extended Data Fig 7d, Supplementary Information). Each structural variant was restricted to a single crypt, except for one deletion that was present in two adjacent crypts that share few mutations, indicating that it occurred during gestation or early childhood.

## Driver mutations

Driver mutations are those that confer a selective advantage during cancer evolution. To search for driver mutations in normal colon, the whole-genome sequences of 571 crypts were supplemented with targeted sequencing of 90 known colorectal cancer genes (Supplementary Table 4) in additional crypts. In total, substitutions in these genes could be evaluated in 1,403 crypts and indels in 1,046. Statistical analysis provided evidence of positive selection on the recessive cancer genes *AXIN2* (three truncating mutations; adjusted $q$ value, 0.004) and *STAG2* (two truncating mutations; adjusted $q$ value, 0.038) indicating that these mutations are probably drivers. Additional mutations that are likely to be drivers were identified in cancer genes with canonical missense hotspot mutations. Nine hotspot mutations in *PIK3CA* (E542K, R38H), *ERBB2* (R678Q, V842I, T862A), *ERBB3* (R475W, R667L) and *FBXW7* (R505C, R658Q) were observed (Extended Data Fig. 8). Given the specificity of these hotspot mutations, most are likely to be drivers. In addition, heterozygous truncating mutations were found in the recessive cancer genes *ARID2*, *ATM* (two), *ATR*, *BRCA2*, *CDK12* (two), *CDKN1B*, *RNF43* (two), *TBL1XR1* and *TP53* (Supplementary Table 5). There was no statistical evidence for selection of truncating mutations in the set of 90 colorectal cancer genes overall. The possibility that some have conferred a growth advantage, however, is not excluded. None of the analysed crypts carried more than one putative driver mutation.

Twenty-three pairs of adjacent crypts shared over 100 SBS1 mutations and thus were probably generated by postnatal fission of crypts. Two pairs carried driver mutations (one nonsense mutation in *AXIN2* and one E542K in *PIK3CA*), although the association of driver mutations with crypt fission is not significant ($P = 0.17$). In one sister crypt the *AXIN2* mutation was rendered homozygous by CN-LOH of chromosome 17q, which demonstrates that clonal evolution is ongoing in normal colon (Figs. 4, 3b).

On the basis of the conservative assumption that just the *AXIN2* and *STAG2* truncating mutations and the missense hotspot mutations in *PIK3CA*, *ERBB2*, *ERBB3* and *FBXW7* are drivers, around 1% of normal colorectal crypts (approximately 150,000 crypts) in a 50–60-year-old individual (the mean age of crypts that were assessed for driver mutations in our cohort was 53 years) carry a driver mutation. As around 40% of people over 70 years old have an adenoma on colonoscopy[40] and

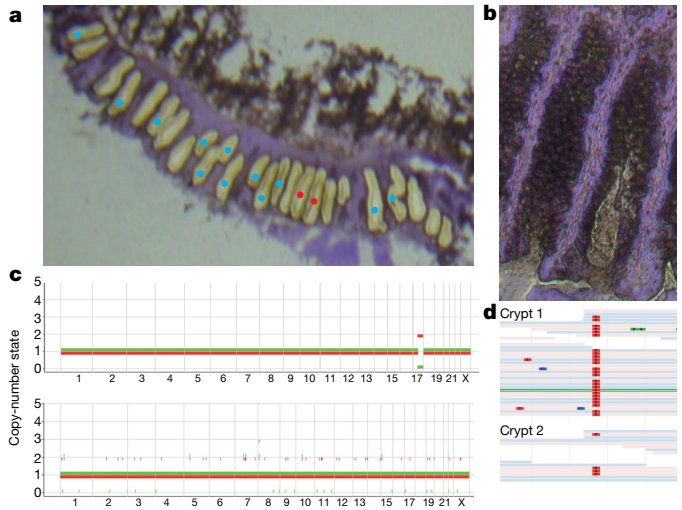



**Fig. 4 | An inactivating mutation in *AXIN2*.** **a**, A section after dissection. Red dots represent crypts with the *AXIN2* mutation; blue dots represent those without it. Crypts without dots were not sequenced successfully. **b**, Crypts with the mutation appeared no different to others. **c**, CN-LOH of one crypt over the *AXIN2* locus. The copy-number state (*y* axis) for every chromosome is shown, with one allele coloured red and the other green. **d**, JBrowse image of reads that support the *AXIN2* mutation. The mutation is red. Of 29 reads, 25 support the mutation in the crypt with CN-LOH; the 4 that do not presumably represent stromal contamination.

around 5% of people develop colorectal cancer over their lifetime[41] (and some of these may arise from more recently acquired driver mutations) only an extremely small proportion of these crypt microneoplasms become a macroscopically detectable adenoma (less than 1 in 375,000) or carcinoma (less than 1 in 3 million) within the following few decades.

## Clonal dynamics of normal epithelium

The distribution of allele fractions of mutations within the crypt informs on the dynamics of turnover of the crypt stem cells. We estimate that the average time since the most recent common ancestor of crypts is 5.5 years (95% confidence interval, 1–10.5), which is similar to previous estimates[14]. Our data are compatible with previous estimates of 7 active stem cells and 1.3 stem cell replacements per year[14], or 5 stem cells and 0.6 stem cell replacements per year[42], but we cannot exclude the possibility of a larger number of stem cells that turn over more frequently (Extended Data Fig. 10, Supplementary Information). The microdissection approach also enabled us to investigate the clonal structure of colonic epithelium beyond the crypt. By comparing the genetic relatedness of crypts with their spatial relatedness, we estimate that crypts undergo fission at a mean rate of once every 27 years (95% confidence interval, 15.9–47.6) (Extended Data Fig. 10, Supplementary Information).

## Comparisons with colorectal cancer

There are marked differences between the genomes of normal colorectal stem cells and those of colorectal cancers. The total mutational burdens of substitutions (10,000–20,000) and indels (1,000–2,000) that are found in most colorectal carcinomas[2] (excluding those with hypermutator phenotypes, in which the burden is usually more than 10-fold higher) are higher than the approximately 3,000 substitutions and 300 indels that are found in most normal crypts from 50–60-year-old individuals (Extended Data Fig. 11a). These differences may be underestimated, as the most recent common ancestor of cancers probably predates that of normal crypts. The high mutational burdens and associated mutational signatures of DNA mismatch repair deficiency and/or mutations in DNA polymerase ε or δ were not found in any normal colorectal crypts, but are present in around 20% of colorectal cancers. Equally striking is the difference between the 0–4 structural changes per normal crypt (with the majority having none; Supplementary Information) and the tens to hundreds per colorectal cancer[43]. In all of these respects, the genomes of normal crypts with driver mutations were similar to those of normal crypts without drivers (Extended Data Fig. 9).

There was no difference in the burden of either sporadic or ubiquitous mutational processes between the crypts of individuals with and without a colorectal cancer (Supplementary Information). If differences in mutational processes in normal cells do underlie why some people develop colon cancer and others do not, these mutational processes must affect only a small proportion of crypts in the colon, or only exert subtle effects on the rate of mutation, such that we could not detect differences between the two groups. The increased base substitution and indel mutation loads in cancers are due to a combination of higher burdens of the ubiquitous mutational signatures found in normal crypts, additional signatures thus far found exclusively in cancers (confirming previous reports[5,44]) and more copy-number changes and structural variation (Extended Data Fig. 11a). The causes of some of these additional mutations in cancer are known (for example, defective mismatch repair and mutations in DNA polymerase ε or δ) but the majority are uncertain.

The relative frequencies of mutated cancer genes differ between colorectal adenomas or carcinomas and normal colorectal cells ($P = 0.003$; Supplementary Information, Extended Data Fig. 11a). Mutations in *APC*, *KRAS* and *TP53* are common in colorectal cancer[37]—accounting for 56% of base-substitution and indel driver mutations (Supplementary Methods)—but comparatively rare among normal crypts with driver mutations (1 in 14). By contrast, mutations in, for example, *ERBB2* and *ERBB3* are common in normal crypts with driver mutations (5 in 14), but rare in colorectal cancer (7 in 631). In the case of *APC* (but not *KRAS*, and perhaps not *TP53*), biallelic inactivation may be required to confer a strong growth advantage; this helps to explain why *APC* may be mutated less frequently in normal colon than *ERBB2* or *ERBB3* (which require a single hit to do so). The results suggest that mutations in *APC*, *KRAS* and *TP53* confer higher likelihoods of conversion to adenoma and carcinoma than mutations in *ERBB2* and *ERBB3*, whereas the latter confer higher likelihoods of crypt colonization by stem cells. There was no detectable difference in the frequency of driver mutations between individuals in our cohort who had colorectal cancer and those who did not (Supplementary Information).

## Discussion

This study has characterized all classes of somatic mutation in hundreds of normal colorectal epithelial stem cells. Our experimental design allows us to gain insights into different facets of the earliest stages of the clonal evolution of colorectal cancers; namely, the range of mutational processes, the frequency of driver mutations and the clonal dynamics of colonic stem cells.

A substantial repertoire of base-substitution and indel mutational processes is operative (some of which are ubiquitous and some sporadic), together with relatively infrequent copy-number changes and genome rearrangements. DNA editing by APOBEC cytidine deaminases occurs in normal colon, albeit only rarely. Many signatures, however, are of unknown aetiology and some appear to be acquired early in life. The presence of five times the age-standard mutational load in all colorectal cells—and potentially many other tissues—in an individual who had undergone chemotherapy provides new insight into the effect of such exposures, and raises questions pertaining to the relationship between mutational load and the relatively modest effect of chemotherapy on cancer risk[45].

Herein, we have revealed the earliest stages of colorectal cancer development. They are characterized by numerous crypts that carry driver mutations, of which only a very small fraction ever manifest as

macroscopic neoplasms. Certain mutated cancer genes appear to foster this pervasive and invisible wave of microneoplastic change, whereas others particularly engender progression to colorectal adenoma and cancer. The conversion of these early microneoplasms to more advanced stages of colorectal neoplasia is associated with the acquisition of increased mutational loads that are composed of base substitutions, indels, structural variants and copy-number changes. More extensive studies of colorectal epithelium will enable the rarer intermediate stages between these early clones and small adenomas to be characterized, and will refine our understanding of the development of the subset of microneoplasms that are more likely to become carcinomas.

The proportion of normal colorectal epithelial cells with driver mutations (1%) is, however, substantially lower than that of other normal tissues so far studied—notably skin (30%)[10] and oesophagus (over 50%)[16]. This may be, at least in part, a consequence of the modular structure of glandular epithelia. The small number of stem cells within a crypt diminishes the probability that a cell with a driver mutation will outcompete its wild-type neighbours. Moreover, even if it does colonize the crypt, a mutant stem cell is entombed in it unless the cell can overcome the largely unknown forces that control clonal expansion through crypt fission. The lower burden of driver mutations in colon relative to endometrium[11,15] (which is also glandular) remains to be investigated.

Fundamental questions are being addressed with respect to differences in the incidence rates of cancer between tissues. The somatic mutational burden in colon and ileum is similar despite the substantially higher incidence of cancer in colon (as previously noted[4]), and therefore does not appear to account for this difference. Whether the total burden of microneoplastic change across the colon and in other tissues more closely correlates with these differences is yet to be determined.

Finally, this study provides a reference perspective on the mutational signatures and driver mutations in normal colon, against which disease states of inflammatory, genetic, neoplastic, degenerative and other aetiologies can be compared. Similar surveys conducted across the range of normal cell types will inform on the universal process of somatic evolution in the human body in health and disease.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1672-7.


1. Fearon, E. R. & Vogelstein, B. A genetic model for colorectal tumorigenesis. *Cell* **61**, 759–767 (1990).
2. Alexandrov, L. B. et al. The repertoire of mutational signatures in human cancer. Preprint at https://www.biorxiv.org/content/10.1101/322859v2 (2019).
3. Sabarinathan, R. et al. The whole genome panorama of cancer drivers. Preprint at https://www.biorxiv.org/content/10.1101/190330v2 (2017).
4. Blokzijl, F. et al. Tissue-specific mutation accumulation in human adult stem cells during life. *Nature* **538**, 260–264 (2016).
5. Roerink, S. F. et al. Intra-tumour diversification in colorectal cancer at the single-cell level. *Nature* **556**, 457–462 (2018).
6. Welch, J. S. et al. The origin and evolution of mutations in acute myeloid leukemia. *Cell* **150**, 264–278 (2012).
7. Bae, T. et al. Different mutational rates and mechanisms in human cells at pregastrulation and neurogenesis. *Science* **359**, 550–555 (2018).
8. Behjati, S. et al. Genome sequencing of normal cells reveals developmental lineages and mutational processes. *Nature* **513**, 422–425 (2014).
9. Lee-Six, H. et al. Population dynamics of normal human blood inferred from somatic mutations. *Nature* **561**, 473–478 (2018).
10. Martincorena, I. et al. High burden and pervasive positive selection of somatic mutations in normal human skin. *Science* **348**, 880–886 (2015).
11. Suda, K. et al. Clonal expansion and diversification of cancer-associated mutations in endometriosis and normal endometrium. *Cell Rep.* **24**, 1777–1789 (2018).
12. Lodato, M. A. et al. Aging and neurodegeneration are associated with increased mutations in single human neurons. *Science* **359**, 555–559 (2018).
13. Hoang, M. L. et al. Genome-wide quantification of rare somatic mutations in normal human tissues using massively parallel sequencing. *Proc. Natl Acad. Sci. USA* **113**, 9846–9851 (2016).
14. Nicholson, A. M. et al. Fixation and spread of somatic mutations in adult human colonic epithelium. *Cell Stem Cell* **22**, 909–918 (2018).
15. Moore, L. et al. The mutational landscape of normal human endometrial epithelium. Preprint at https://www.biorxiv.org/content/10.1101/505685v1 (2018).
16. Martincorena, I. et al. Somatic mutant clones colonize the human esophagus with age. *Science* **362**, 911–917 (2018).
17. Jaiswal, S. et al. Age-related clonal hematopoiesis associated with adverse outcomes. *N. Engl. J. Med.* **371**, 2488–2498 (2014).
18. Xie, M. et al. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat. Med.* **20**, 1472–1478 (2014).
19. McKerrell, T. et al. Leukemia-associated somatic mutations drive distinct patterns of age-related clonal hemopoiesis. *Cell Rep.* **10**, 1239–1245 (2015).
20. Genovese, G. et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N. Engl. J. Med.* **371**, 2477–2487 (2014).
21. Potten, C. S., Kellett, M., Roberts, S. A., Rew, D. A. & Wilson, G. D. Measurement of in vivo proliferation in human colorectal mucosa using bromodeoxyuridine. *Gut* **33**, 71–78 (1992).
22. Cheng, H. & Leblond, C. P. Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. V. Unitarian Theory of the origin of the four epithelial cell types. *Am. J. Anat.* **141**, 537–561 (1974).
23. Lopez-Garcia, C., Klein, A. M., Simons, B. D. & Winton, D. J. Intestinal stem cell replacement follows a pattern of neutral drift. *Science* **330**, 822–825 (2010).
24. Snippert, H. J. et al. Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell* **143**, 134–144 (2010).
25. Griffiths, D. F., Davies, S. J., Williams, D., Williams, G. T. & Williams, E. D. Demonstration of somatic mutation and colonic crypt clonality by X-linked enzyme histochemistry. *Nature* **333**, 461–463 (1988).
26. Winton, D. J. & Ponder, B. A. J. Stem-cell organization in mouse small intestine. *Proc. R. Soc. B* **241**, 13–18 (1990).
27. Kozar, S. et al. Continuous clonal labeling reveals small numbers of functional stem cells in intestinal crypts and adenomas. *Cell Stem Cell* **13**, 626–633 (2013).
28. Barker, N. et al. Crypt stem cells as the cells-of-origin of intestinal cancer. *Nature* **457**, 608–611 (2009).
29. Rouhani, F. J. et al. Mutational history of a human cell lineage from somatic to induced pluripotent stem cells. *PLoS Genet.* **12**, e1005932 (2016).
30. Viel, A. et al. A specific mutational signature associated with DNA 8-Oxoguanine persistence in MUTYH-defective colorectal cancer. *EBioMedicine* **20**, 39–49 (2017).
31. Alexandrov, L. B. et al. Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
32. Nik-Zainal, S. et al. Mutational processes molding the genomes of 21 breast cancers. *Cell* **149**, 979–993 (2012).
33. Chan, K. et al. An APOBEC3A hypermutation signature is distinguishable from the signature of background mutagenesis by APOBEC3B in human cancers. *Nat. Genet.* **47**, 1067–1072 (2015).
34. Boot, A. et al. Identification of novel mutational signatures in Asian oral squamous cell carcinomas associated with bacterial infections. Preprint at https://www.biorxiv.org/content/10.1101/368753v3 (2019).
35. Haradhvala, N. J. et al. Mutational strand asymmetries in cancer genomes reveal mechanisms of DNA damage and repair. *Cell* **164**, 538–549 (2016).
36. Wolf, J. et al. Peripheral blood mononuclear cells of a patient with advanced Hodgkin's lymphoma give rise to permanently growing Hodgkin–Reed Sternberg cells. *Blood* **87**, 3418–3428 (1996).
37. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330–337 (2012).
38. Bomme, L. et al. Cytogenetic analysis of colorectal adenomas: karyotypic comparisons of synchronous tumors. *Cancer Genet. Cytogenet.* **106**, 66–71 (1998).
39. Andersen, C. L. et al. Frequent occurrence of uniparental disomy in colorectal cancer. *Carcinogenesis* **28**, 38–48 (2007).
40. Corley, D. A. et al. Variation of adenoma prevalence by age, sex, race, and colon location in a large population: implications for screening and quality programs. *Clin. Gastroenterol. Hepatol.* **11**, 172–180 (2013).
41. Cancer Research UK. *Bowel Cancer Incidence Statistics* https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/incidence#heading-Seven (accessed August 2018).
42. Stamp, C. et al. Predominant asymmetrical stem cell fate outcome limits the rate of niche succession in human colonic crypts. *EBioMedicine* **31**, 166–173 (2018).
43. Li, Y. et al. Patterns of structural variation in human cancer. Preprint at https://www.biorxiv.org/content/10.1101/181339v1 (2017).
44. Lugli, N. et al. Enhanced rate of acquisition of point mutations in mouse intestinal adenomas compared to normal tissue. *Cell Rep.* **19**, 2185–2192 (2017).
45. Travis, L. B. Therapy-associated solid tumors. *Acta Oncol.* **41**, 323–333 (2002).

# Methods

### Data reporting

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

### Human tissues

We obtained healthy colonic biopsies from four cohorts (Supplementary Table 1). The first represents seven deceased organ donors ranging in age from 36 to 67, from whom colonic and small-intestinal biopsies were taken at the time of organ donation (REC 15/EE/0152). The second represents individuals aged 60 to 72 who were having a colonoscopy after a positive faecal occult blood test as part of the bowel cancer screening programme (Ethical approval 08-H0308-13); we selected 16 individuals who were not found to have either an adenoma or a carcinoma on colonoscopy, and 15 who were found to have a colorectal carcinoma (the normal biopsies that we used were distant from these lesions). The third cohort represents three paediatric patients who underwent routine colonoscopy to exclude inflammatory bowel disease and who were found to have a completely normal intestinal mucosa macroscopically and histologically (REC 12/EE/0482). The final cohort comprised one 78-year-old man with oesophageal cancer who underwent a rapid autopsy (REC 13/EE/0043). This individual had been treated with palliative chemotherapy of epirubicin, oxaliplatin and capecitabine within the three months before death; given that monoclonal conversion within crypts takes place on the scale of years, mutations caused by these chemotherapies are likely to be restricted to a small proportion of stem cells per crypt and so unlikely to be detected. All samples were obtained with informed consent and the studies were approved by East of England Research Ethics Committees.

### Laser-capture microdissection of colonic crypts

Fresh frozen biopsies were embedded in optimal cutting temperature compound. Sections of 30 μm were fixed in methanol for 5 min, washed three times with phosphate-buffered saline and stained with Gill's haematoxylin for 20 s. Crypts were isolated by laser-capture microdissection, and collected in separate wells of a 96-well plate. They were lysed using the Arcturus PicoPure Kit (Applied Biosystems) according to the manufacturer's instructions. DNA library preparation then proceeded without clean-up or quantification.

### Library preparation

Two library-preparation methods were used for laser-capture microdissected material: in initial experiments, sonication was used to fragment DNA; and later, an enzymatic fragmentation method was implemented as it could make libraries from even lower input. Comparison of the two methods showed no difference in mutation calls once post-processing filters (described below) had been implemented. All samples in this study were processed using an Agilent Bravo Workstation (Option B; Agilent Technologies).

For sonication libraries, lysate from laser-capture microdissection (20 μl) was mixed with 100 μl TE buffer 1(0 mM Tris-HCl, 1 mM EDTA) (Ambion) and DNA was fragmented using focused acoustics (Covaris LE220; Covaris, Inc.). Fragmented DNA was mixed with 80 μl Ampure XP beads (Beckman Coulter). After a 5-min binding reaction and magnetic bead separation, genomic DNA was washed twice with 75% ethanol. Beads were resuspended in 20 μl nuclease-free water (Ambion) and processed immediately for DNA library construction. Each sample (20 μl) was mixed with 2.8 μl NEBNext Ultra II End Prep Reaction Buffer and 1.25 μl NEBNext Ultra II End Prep Enzyme Mix (New England BioLabs), and incubated on a thermal cycler for 30 min at 20 °C then 30 min at 65 °C. Following DNA fragmentation and A-tailing, each sample was incubated for 20 min at 20 °C with a mixture of 30 μl ligation mix and 1 μl ligation enhancer (New England BioLabs), 0.9 μl nuclease-free water (Ambion) and 0.1 μl duplexed adapters (100 μM; 5′-ACACTCTTTCCCTACACGACGCTCTTCCGATC*T-3′, 5′-phos-GATCGGAAGAGCGGTTCAGCAGGAATGCCGAG-3′). Adapter-ligated libraries were purified using Ampure XP beads by addition of 65 μl Ampure XP solution (Beckman Coulter) and 65 μl TE buffer (Ambion). After elution and bead separation, DNA libraries (21.5 μl) were amplified by PCR by addition of 25 μl KAPA HiFi HotStart ReadyMix (KAPA Biosystems), 1 μl PE1.0 primer (100 μM; 5′-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATC*T-3′) and 2.5 μl iPCR-Tag (40 μM; 5′-CAAGCAGAAGACGGCATACGAGATXGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC-3′), in which 'X' represents one of 96 unique 8-base indexes The sample was then mixed and thermal-cycled as follows: 98 °C for 5 min, then 12 cycles of 98 °C for 30 s, 65 °C for 30 s, 72 °C for 1 min and finally 72 °C for 5 min. Amplified libraries were purified using a 0.7:1 volumetric ratio of Ampure Beads (Beckman Coulter) to PCR product and eluted into 25 μl nuclease-free water (Ambion). DNA libraries were adjusted to 2.4 nM and sequenced on the HiSeq X platform (Illumina) according to the manufacturer's instructions, with the exception that we used iPCR-Tag (5′-AAGAGCGGTTCAGCAGGAATGCCGAGACCGATCTC-3′) to read the library index.

For enzymatic fragmentation, lysate from laser-capture microdissection (20 μl) was mixed with 50 μl Ampure XP beads (Beckman Coulter) and 50 μl TE buffer (10 mM Tris-HCl, 1 mM EDTA) (Ambion) at room temperature. After a 5-min binding reaction and magnetic bead separation, genomic DNA was washed twice with 75% ethanol. Beads were resuspended in 26 μl TE buffer and the bead–genomic DNA slurry was processed immediately for DNA library construction. Each sample (26 μl) was mixed with 7 μl 5X Ultra II FS buffer and 2 μl Ultra II FS enzyme (New England BioLabs), and incubated on a thermal cycler for 12 min at 37 °C then 30 min at 65 °C. Following DNA fragmentation and A-tailing, each sample was incubated for 20 min at 20 °C with a mixture of 30 μl ligation mix and 1 μl ligation enhancer (New England BioLabs), 0.9 μl nuclease-free water (Ambion) and 0.1 μl duplexed adapters (100 μM; 5′-ACACTCTTTCCCTACACGACGCTCTTCCGATC*T-3′, 5′-phos-GATCGGAAGAGCGGTTCAGCAGGAATGCCGAG-3′). Adapter-ligated libraries were purified using Ampure XP beads by addition of 65 μl Ampure XP solution (Beckman Coulter) and 65 μl TE buffer (Ambion). After elution and bead separation, DNA libraries (21.5 μl) were amplified by PCR by addition of 25 μl KAPA HiFi HotStart ReadyMix (KAPA Biosystems), 1 μl PE1.0 primer (100 μM; 5′-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATC*T-3′) and 2.5 μl iPCR-Tag (40 μM; 5′-CAAGCAGAAGACGGCATACGAGATXGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC-3′), in which 'X' represents one of 96 unique 8-base indexes. The sample was then mixed and thermal-cycled as follows: 98 °C for 5 min, then 12 cycles of 98 °C for 30 s, 65 °C for 30 s, 72 °C for 1 min and finally 72 °C for 5 min. Amplified libraries were purified using a 0.7:1 volumetric ratio of Ampure Beads (Beckman Coulter) to PCR product and eluted into 25 μl nuclease-free water (Ambion). DNA libraries were adjusted to 2.4 nM and sequenced on the HiSeq X platform (Illumina) according to the manufacturer's instructions, with the exception that we used iPCR-Tag (5′-AAGAGCGGTTCAGCAGGAATGCCGAGACCGATCTC-3′) to read the library index.

### Whole-genome sequencing

We generated paired-end sequencing reads (150 bp) using Illumina XTEN® machines, resulting in a coverage of around 15× per sample. In 94% of the whole-genome-sequenced crypts that were included for statistical analysis, over 90% of the callable genome was covered by more than 10 reads. Sequences were aligned to the human reference genome (NCBI build37) using BWA-MEM.

### Targeted sequencing

A 2.3-Mb capture panel was designed in-house to pull down genes that are known or suspected to play a role in neoplasia. We performed custom RNA bait design following the manufacturer's guidelines (SureSelect;

Agilent). Samples were multiplexed on flow cells and subjected to paired-end sequencing (75-bp reads) using Illumina HiSeq2000 machines. One 96-well plate of samples was sequenced on each lane, but as tissue recovery was variable, a range of coverage was achieved. Sequences were aligned to the human reference genome (NCBI build37) using BWA-align.

### Calling substitutions

Substitution calling was broken down into three steps: mutation discovery; filtering to produce a list of clean sites; and genotyping, in which the presence or absence of every mutation in every sample is evaluated.

First, mutations were discovered using the cancer variants through expectation maximization (CaVEMan) algorithm[46]. CaVEMan uses a naive Bayesian classifier to derive the probability of all possible genotypes at each nucleotide. CaVEMan copy-number options were set to major copy number 5 and minor copy number 2 for normal clones, as in our experience this maximizes sensitivity. The algorithm was run using an unmatched normal to be able to derive phylogenies: had another sample from the same individual been treated as a matched normal, early embryonic mutations would have been treated as germline and discarded, resulting in incorrect trees.

Second, a number of post-processing filters were applied. These included filtering against a panel of 75 unmatched normal samples to remove common single-nucleotide polymorphisms (SNPs), post-processing as described previously[32] and two filters (only applied to whole-genome-sequencing data) designed to remove mapping artefacts associated with BWA-MEM: the median alignment score of reads supporting a mutation should be greater than or equal to 140, and fewer than half of these reads should be clipped. The library-preparation protocol for microbiopsies produced shorter library insert sizes than standard methods. Reads could therefore overlap, resulting in double counting of mutant reads. Fragment-based statistics were generated to prevent the calling of variants that were supported by a low number of fragments. Variants were annotated by ANNOVAR[47] and fragment-based statistics (fragment coverage, number of fragments supporting the variant, fragment-based allele fraction) were calculated for each variant after the exclusion of marked PCR duplicates. In the rare event of discordance in the called base at the variant position between overlapping paired-end reads, the base with the highest-quality score was selected. Fragment-based statistics were calculated separately for high-quality fragments (alignment scores greater than or equal to 40 and base scores greater than or equal to 30). Variants that were supported by at least three-high quality fragments were retained and used for the next stage of variant filtering. Inspection of variants specific to laser-capture microdissection experiments revealed that the vast majority were present within inverted repeats capable of forming hairpin structures, that they were supported by reads with very similar alignment start position (and so not marked as PCR duplicates) and were primarily located close to the alignment start within the supporting reads. These variants commonly coincided with other proximal variants (1–30 bp), but filtering based on variant proximity would also remove actual kataegis events. In silico modelling of the potential hairpin showed that the variants were aligning to each other in the stem of the structure, but could not form a base pair, whereas all other bases could. The artefacts are probably the consequence of erroneous processing of cruciform DNA (either existing before DNA isolation or formed during library preparation) by the enzymatic digestion protocol applied. We considered modelling the hairpin structures to filter these variants, but given the fact that read clustering (i.e., similar alignment position) serves as a hallmark for these artefacts, we opted to use the proximity of the variant to the alignment start, and the standard deviation (s.d.) and median absolute deviation (MAD) of the variant position within the supporting reads, as features for filtering. These statistics were calculated separately for reads aligned to positive and negative strands. In cases in which the variant was supported by a low number of reads (i.e., 0–1 reads) for one of the strands, the filtering was based only on the statistics generated for the other strand. Per variant, if one of the strands had too few supporting reads, it was required for the other strand that either: i) there should be 90% or more supporting reads to report the variant within the first 15% of the read starting from the alignment start; or ii) the statistics MAD > 0 and s.d. > 4. Per variant, if both strands were supported by sufficient reads it was required for both strands separately that: i) there should be 90% or more supporting reads to report the variant within the first 15% of the read; ii) the statistics MAD > 2 and s.d. > 2; or iii) the other strand should have the statistics MAD > 1 and s.d. > 10 (i.e., the variant is retained if the other strand demonstrates strong measures of variance). In our experience, the proposed strategy greatly reduces the number of artefactual variants while retaining all other variants—as assessed by running the last filtering step on whole-genome-sequencing data from experiments that were not laser-capture microdissections.

Third, mutations were genotyped in every sample. A pile-up of all the samples from a given individual was constructed, in which the numbers of mutant and wild-type reads in every sample over every site that had been called in any sample from that person were counted. Only reads with a mapping quality of 30 or above and bases with a base quality of 30 or above were counted. After applying these filters, mutations were genotyped on the basis of the number of mutant and wild-type reads at each locus. Mutations were called on the basis of a VAF greater than 0.2, a depth greater than 7 and at least 4 mutant reads. If the depth over a locus was less than 7 in a given sample, or if there was more than one mutant read but the other criteria were not met, the genotype was set to NA (not applicable) for tree-construction purposes. Loci that were set to NA in more than one third of the samples were removed for construction of the phylogeny. Positions were called as germline if they were either called as present or NA in all of the samples from a given individual.

Around 1.2% of all mutations were present in the coding regions of the genome. All mutations in coding regions are provided (Supplementary Table 3).

### Calling short indels

As for substitutions, calling of indels was broken down into mutation discovery, filtering and genotyping. Mutations were called with the Pindel algorithm[48] using an unmatched normal. Post-processing filters were applied as described previously[32], and the number of mutant and wild-type reads was tabulated as above. The same dataset-specific filters were applied as for substitutions. Indels were then genotyped on the basis of a VAF greater than 0.2, a depth of at least 10 and support of at least 5 mutant reads.

### Calling structural variants

Genomic rearrangements were called using the BRASS algorithm (https://github.com/cancerit/BRASS). Abnormally paired read pairs from whole-genome sequencing were grouped and filtered by read remapping. Read-pair clusters for which more than 50% of the reads mapped to microbial sequences were removed, as were rearrangements for which the breakpoint could not be reassembled. Candidate breakpoints were matched to copy-number breakpoints defined by ASCAT (see below) within 10 kb. Only structural variants in which the two breakpoints were more than 1,000 base pairs apart were considered. Structural variants were called against a matched normal skin or blood sample when available and against another crypt from the same individual with good coverage when not.

### Calling copy number

Copy-number changes were called using the allele-specific copy number analysis of tumours (ASCAT) algorithm[49]. The same matched normal sample was used as for calling structural variants. For additional validation of copy-number changes in normal colon, the QDNAseq algorithm[50] was run. ASCAT uses both the read depth and ratios of heterozygous SNPs to determine an allele-specific copy number, whereas QDNAseq relies solely on variations in sequencing coverage. To call amplifications

and deletions in the colonic microbiopsy cohort, only those that both were called by ASCAT and showed a clear departure from the background $\log_2$-transformed ratio by QDNAseq were retained. To call CN-LOH in this cohort, all such events called by ASCAT were checked visually on JBrowse[51] to verify an imbalance of parental SNPs. Only crypts with a coverage of more than 10×, for which copy-number changes could be reliably detected, were used.

### Detection of driver variants and positive selection

Driver mutations were detected through both an unbiased dN/dS method and manual annotation. For these analyses, the CaVEMan and Pindel calls were used without post-processing filters (such as requiring a VAF cut-off of greater than 0.2) to maximize our sensitivity. All putative driver variants were visually inspected using JBrowse[51], thus we could afford a higher false-positive rate in the mutation discovery phase.

The statistical model dNdScv[52] was used to conduct three tests: first, using only the whole-genome-sequencing data, an analysis of selection over all genes; second, using combined whole-genome and targeted-sequencing data, over all the genes covered by the bait set; and finally, again using this combined dataset, over 90 selected cancer genes (Supplementary Table 4, Supplementary Information).

Manual annotation of driver variants on the basis of previous knowledge complemented this. A list of 90 colorectal cancer genes (appendix) curated from the literature that were also covered by the bait set were intersected with the list of substitutions and indels from combined whole-genome and targeted sequencing. Mutations were annotated as putative drivers either if they were missense mutations that fell in an oncogene hotspot (on the basis of visualization of the distribution of mutations in the gene on COSMIC[53]), or if they were truncating mutations that fell in a tumour suppressor gene.

Structural variants that might act as drivers were assessed by intersection of genes involved in each structural variant with the twelve genes involved in gene fusions that have been reported in colorectal cancer in COSMIC (*VTI1A*, *TCF7L2*, *TPM3*, *NTRK1*, *PTPRK*, *RSPO3*, *ETV6*, *NTRK3*, *EIF3E*, *RSPO2*, *C2orf44* and *ALK*). No fusion genes were found. None of the genes involved in structural variants in our data overlapped with the list of 90 cancer genes used for assessing substitutions and indels, and there were no genes that were affected by more than one structural variant. No high-level copy-number amplifications were observed and there were no homozygous deletions.

Note that the frequency of driver mutations is low and as such we cannot estimate a per-gene driver frequency. All we can do to derive a meaningful estimate is to pool our driver mutations. In addition, coverage may fluctuate even within a gene. Some portions of a gene may be well-covered in 1,000 crypts, and others in 2,000 crypts. The approach that we took was to calculate, for the average exonic base pair in our 90 cancer genes, the number of crypts in which that base pair was covered by 8 or more reads (for substitutions) and by 10 or more reads (for indels). Of all bases in the targeted panel across all crypts, 64% are covered by 8 or more reads, which equates to a number of callable bases equivalent to having sequenced about 1,400 crypts with perfect coverage over every base in every crypt. This average number of crypts in which all base pairs achieve good coverage becomes the denominator for calculating the frequency of driver mutations (with the number of drivers observed in the dataset as the numerator). A similar approach can be taken with indels. Our estimate of 1% uses a global correction, on the assumption that under-representation and over-representation will even itself out when estimating the total frequency of driver mutations in the whole dataset.

### Estimation of the frequency of driver mutations in cancer

Publicly available colorectal cancer mutation calls were obtained from The Cancer Genome Atlas (TCGA) network[37]. Driver mutations were annotated manually in the same way as in our dataset: only mutations that fell in the 90 genes that we had selected were considered, and they were annotated as putative drivers either if they were missense mutations that fell in an oncogene hotspot (on the basis of visualization of the distribution of mutations in the gene on COSMIC[53]), or if they were truncating mutations that fell in a tumour suppressor gene.

### Construction of phylogenies

Phylogenies are used in this analysis for timing mutations. The most informative branches in this case are the long branches shared by a small number of crypts, which are very robust to all methods of tree construction. Trees were built using maximum parsimony, with substitutions called as described above. For every individual, the input matrix of mutation calls was bootstrapped 100 times. Phylogenies were constructed for each replicate using the Wagner method of the Mix programme from the PHYLIP suite of tools[54]. The consensus phylogeny was constructed from 100 bootstrap runs using the extended majority rule method for the Consense programme from the PHYLIP suite of tools[54].

Across all phylogenies, a mean of 10% and a median of 1.5% of mutations per tree did not fit the trees perfectly. Phylogenies with more crypts had more mutations that fitted imperfectly. Consider a mutation that is actually present in 50 crypts. Even with 15× coverage over the site in every sample, and with every crypt completely clonal, if we simulate resampling of mutant reads from the binomial distribution (with size of 15 and probability of 0.5), 17% of the time the mutation will have fewer than the 3 reads required to call it in at least one sample. Variation in sequencing depth, clonality and sequencing errors would further decrease the probability of calling the mutation perfectly in every sample. Nodes across all our phylogenies had mean bootstrapping values of 0.77 and median bootstrapping values of 0.99. Branches at the very top of the phylogenies, which probably represent embryonic cell divisions, are supported by only a few mutations and have lower support because in a given bootstrap sample the couple of mutations that support this node may be omitted. Longer shared branches almost always have bootstrapping values of 1. These longer shared branches are those that are most important to our analyses, because they are the most informative when timing mutational signatures relative to one another and because they represent postnatal crypt fission events. To further increase our confidence in our phylogenies, we validated them by reconstructing them with indels. To do this, the same procedure as for substitutions was followed for indel matrices. As there were fewer indels than substitutions, nodes in indel phylogenies were generally reconstructed with lower confidence than in substitution phylogenies, but they broadly agree. Of the nodes reconstructed with 90% or greater confidence in the indel tree, 85% were present with exactly the same set of descendants in the substitution trees. Any errors in the phylogenies should be relatively minor and not affect our downstream analyses.

The program that was used for inferring phylogeny provided the topology of the tree but not the assignment of mutations. Mutations from the input matrix of genotypes therefore have to be reassigned to branches. To assign a set of mutation calls with no false negative and no false positives to a tree, each branch of the tree was considered in turn. If a mutation was called in all the descendants of a given branch, and in no samples that were not descendants of the branch, mutations were assigned to that branch.

Some colonic microbiopsies had low coverage and stromal contamination. For this reason, we did not expect mutations to fit the tree perfectly, as a mutation that was truly present in a colony might be missed if too few supporting reads are found. Mutations were only assigned to the tree to determine the mutational processes active at a particular time. We reasoned that it was preferable to assign only mutations that fit the tree perfectly and adjust the branch lengths based on the power to call mutations at a given branch, rather than attempting to assign mutations that fit the tree imperfectly. Using the clonality and coverage of all descendants of a branch, the proportion of true substitutions or indels on the branch that would be first discovered (whether by CaVEMan or Pindel) and then genotyped as present according to the criteria

described above was calculated. The observed branch length was then adjusted by dividing by this proportion. Adjustment proportions can be found in Supplementary Table 7. This was done for both substitutions and indels, but not for structural variants and for larger copy-number changes, owing to a lack of data: most branches have no large variants and so could not be extended appropriately. Rearrangements and copy-number changes were assigned to phylogenies manually.

### Extraction of mutational signatures

Mutational signatures were extracted using the mutations assigned to every branch of a phylogeny as a 'sample'. This allows better discrimination of mutational processes that may occur at different times within the same cell. Mutations were categorized following the method used by the Mutational Signatures working group of the PCAWG[2]. Single-base substitutions were categorized into 96 classes according to the identity of the pyrimidine-mutated base pair, and the base 5′ and 3′ to it. Doublet-base substitutions were categorized into 78 classes according to the identity of the reference and alternative bases. Indels were classified into 83 classes according to whether they were an insertion or a deletion, the identity of the inserted or deleted base, the length of the mononucleotide tract in which they occurred and the degree of homology with the surrounding sequence (Fig. 1a).

Signatures were extracted using a hierarchical Dirichlet process (HDP)[55] (https://github.com/nicolaroberts/hdp). Code and the input mutations are provided at https://github.com/HLee-Six/colon_microbiopsies. First, the algorithm was conditioned on the set of mutational signatures that have found to be operative in colorectal cancers in PCAWG[2]: SBS1, SBS2, SBS3, SBS5, SBS13, SBS16, SBS17a, SBS17b, SBS18, SBS25 (included although it is not found in colorectal cancer because the similarity with the mutational profile with crypts from one individual had been previously noted), SBS28, SBS30, SBS37, SBS40, SBS41, SBS43, SBS45, SBS49, DBS, DBS3, DBS4, DBS6, DBS7, DBS8, DBS9, DBS10, DBS11, ID1, ID2, ID3, ID4, ID5, ID6, ID7, ID8, ID10 and ID14. This allows simultaneous discovery of new signatures and matching to known ones. Nine SBS, two DBS and five ID signatures were discovered (Extended Data Fig. 2). Despite pre-conditioning, signatures that were perfectly correlated in all samples were still amalgamated. This occurred, for example, with SBS1, SBS5 and SBS18. Therefore, expectation maximization was used to deconvolute all HDP signatures into known PCAWG signatures. If a signature that was reconstituted from the components that expectation maximization extracted (only including PCAWG signatures that accounted for at least 10% of mutations in each sample to avoid overfitting) had a cosine similarity to the HDP signature of more than 0.95, the signature was presented as its expectation maximization deconvolution. Three HDP signatures met these criteria: the HDP SBS1 signature was deconvoluted into a mixture of PCAWG SBS1, PCAWG SBS5 and PCAWG SBS18; the HDP DBSA signature was deconvoluted into PCAWG DBS2, PCAWG DBS4, PCAWG DBS6, PCAWG DBS9 and PCAWG DBS11; and the HDP IDC signature was deconvoluted into PCAWG ID1, PCAWG ID2 and PCAWG ID5 (Extended Data Fig. 3). To test the robustness of this signature analysis, other signature-extraction methods were used: HDP with no pre-conditioning, the non-negative matrix factorization (NMF) method used in a previous study[4] and a version of the NMF algorithm used by in another study[2]. These all produced comparable results (Extended Data Fig. 4).

### Timing SBSA and SBSB throughout life

Five patients had informative clades with branch points that allowed us to time SBSA. Plotting the cumulative amount of SBSA versus SBS1 at each node in these clades (Extended Data Fig. 6aq), we observed that for each, the rate of accumulation of SBSA relative to SBS1 was high in early branch points and then slowed down almost to zero on all branches but for one (a branch of patient ao; for this patient, SBSA continued to be acquired, albeit at a slow rate).

We can take the inflexion point on the graph of cumulative SBSA versus SBS1 to be the upper limit of the point in time when SBSA slowed down. This provides an upper bound for three reasons. First, when we observe the presence of a signature on a branch, we know that the causative process must have been active at some point during the lifetime of the branch, but we cannot say when on the branch it occurs; it might have ended long before the branch did. Second, if the time to the most recent common ancestor of the crypt is longer than 0, the age at which this stopped would be earlier. Third, if the SBS1 mutation rate is increased in early life—as it may be during the rapid growth of the embryo—the age at which the inflexion point occurs would be earlier.

Using these five informative clades, and assuming a clock-like but personalized rate of SBS1 accumulation (i.e., each patient can accumulate SBS1 at their own constant rate), we found that the upper bound of the age at which SBSA slowed was: 9.7 years (patient h); 7.1 years (patient z); 2.4 years (patient am); 20.1 years (patient aa); and 9.6 years (patient ao). There are no branches that begin after 10 years of age with a high ratio of SBSA to SBS1.

The most informative branch point is the earliest inflexion point; the estimate of 2.4 years from patient aa is therefore, perhaps, our best estimate. Nonetheless, we did not want to base our statement on a single patient, and so 10 years was given in the text as 4 patients had branches that ended before 10 years of age.

A similar argument can be made for SBSB (Extended Data Fig. 6aq). For SBSB, however, only two clades were informative. The estimated upper bounds of age for SBSB activity were 2.4 years old (this was the same inflexion point as for SBSA in patient aa) and 6.4 years old (patient ai). For patient ai, a reasonable amount of SBSB is still acquired after this branch point. If the ratio of accumulation of SBSB versus SBS1 continued at the same rate as before this branch point, the number of SBSB mutations seen in the terminal branches would have been observed by 8.2 years of age.

### Analysis of telomere length

Telomere length was estimated from whole-genome-sequencing data using Telomerecat. Telomerecat is a Python-based software package for estimating telomere length from short-read whole-genome-sequencing data[56]. It functions by classifying paired-end reads as either fully or partially telomeric on the basis of the canonical hexamer TTAGGG, and uses that ratio to estimate an average telomere length. Notably, Telomerecat measures telomeres and also accounts for interstitial telomeric repeats. It is ploidy and species agnostic (assuming that the telomere hexamer is the canonical mammalian signature of $TTAGGG_n$). Telomerecat has four main stages: 1) identification of all telomeric or partially telomeric read pairs and creation of a subsetted bam file that contains only these reads; 2) classification of telomeric read-pairs as intratelomeric, boundary, junction-spanning or intrachromosomal; 3) error correction of boundary or junction-spanning read pairs; and 4) estimation of telomere length based on the ratio of intratelomeric and boundary or junction-spanning read pairs.

Telomerecat has been validated on whole-genome DNA-sequencing files from both tumour and normal samples[56]. Its results show concordance with an established method of telomere length measurement, the mean telomere restriction fragment (mTRF) technique. Alternative packages are available, notably Computel[57], Telseq[58] and TelomereHunter[59]. They all have respective strengths and have been benchmarked through their methods publications. We have opted for Telomerecat as it provides an estimate of telomere length at base-pair resolution, while also correcting for variations in sequencing depth in a ploidy-agnostic manner.

We ran Telomerecat on 445 crypt bam files with good coverage and clonality to generate estimates of telomere length. Telomerecat was threaded across 10 cores and 100 simulation cycles were requested per run. The values displayed are the median telomere length across all chromosomes in that sample, measured in base pairs.

### Statistical analyses

All statistical analyses were performed in R (Supplementary Information). Code can be found at https://github.com/HLee-Six/colon_microbiopsies.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

Whole-genome and targeted sequencing data are deposited in the European Genome-phenome Archive (EGA) with accession codes EGAD00001004192 and EGAD00001004193. Images of microdissections and the physical distances between crypts are available on Mendeley Data (https://data.mendeley.com/datasets/zv6xrjxftw/1) by searching for the title of this article. All other data are available from the authors on request.

## Code availability

Code for statistical analyses is provided as part of the Supplementary Information. Custom R scripts and their input data for signature analysis are available on GitHub at https://github.com/HLee-Six/colon_microbiopsies. All other code is available from the authors on request.


46. Jones, D. et al. cgpCaVEManWrapper: Simple execution of CaVEMan in order to detect somatic single nucleotide variants in NGS data. *Curr. Protoc. Bioinformatics* **56**, 15.10.1–15.10.18 (2016).
47. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
48. Raine, K. M. et al. cgpPindel: Identifying somatically acquired insertion and deletion events from paired end sequencing. *Curr. Protoc. Bioinformatics* **52**, 15.7.1–15.7.12 (2015).
49. Van Loo, P. et al. Allele-specific copy number analysis of tumors. *Proc. Natl Acad. Sci. USA* **107**, 16910–16915 (2010).
50. Scheinin, I. et al. DNA copy number analysis of fresh and formalin-fixed specimens by shallow whole-genome sequencing with identification and exclusion of problematic regions in the genome assembly. *Genome Res.* **24**, 2022–2032 (2014).
51. Buels, R. et al. JBrowse: a dynamic web platform for genome visualization and analysis. *Genome Biol.* **17**, 66 (2016).
52. Martincorena, I. et al. Universal patterns of selection in cancer and somatic tissues. *Cell* **171**, 1029–1041 (2017).
53. Forbes, S. A. et al. COSMIC: somatic cancer genetics at high-resolution. *Nucleic Acids Res.* **45**, D777–D783 (2017).
54. Felsenstein, J. PHYLIP—Phylogeny inference package (version 3.2).*Cladistics* **5**, 164–166 (1989).
55. Roberts, N. *Patterns of somatic genome rearrangement in human cancer. PhD thesis, Univ. Cambridge* (Wellcome Trust Sanger Institute, 2018).
56. Farmery, J. H. R., Smith, M. L., Bioresource, N., Diseases, R. & Lynch, A. G. Telomerecat: a ploidy-agnostic method for estimating telomere length from whole genome sequencing data. *Sci. Rep.* **8**, 1300 (2018).
57. Nersisyan, L. & Arakelyan, A. Computel: computation of mean telomere length from whole-genome next-generation sequencing data. *PLoS One* **10**, e0125201 (2015).
58. Ding, Z. et al. Estimating telomere length from whole genome sequence data. *Nucleic Acids Res.* **42**, e75 (2014).
59. Feuerbach, L. et al. TelomereHunter: telomere content estimation and characterization from whole genome sequencing data. Preprint at https://www.biorxiv.org/content/10.1101/065532v1 (2016).

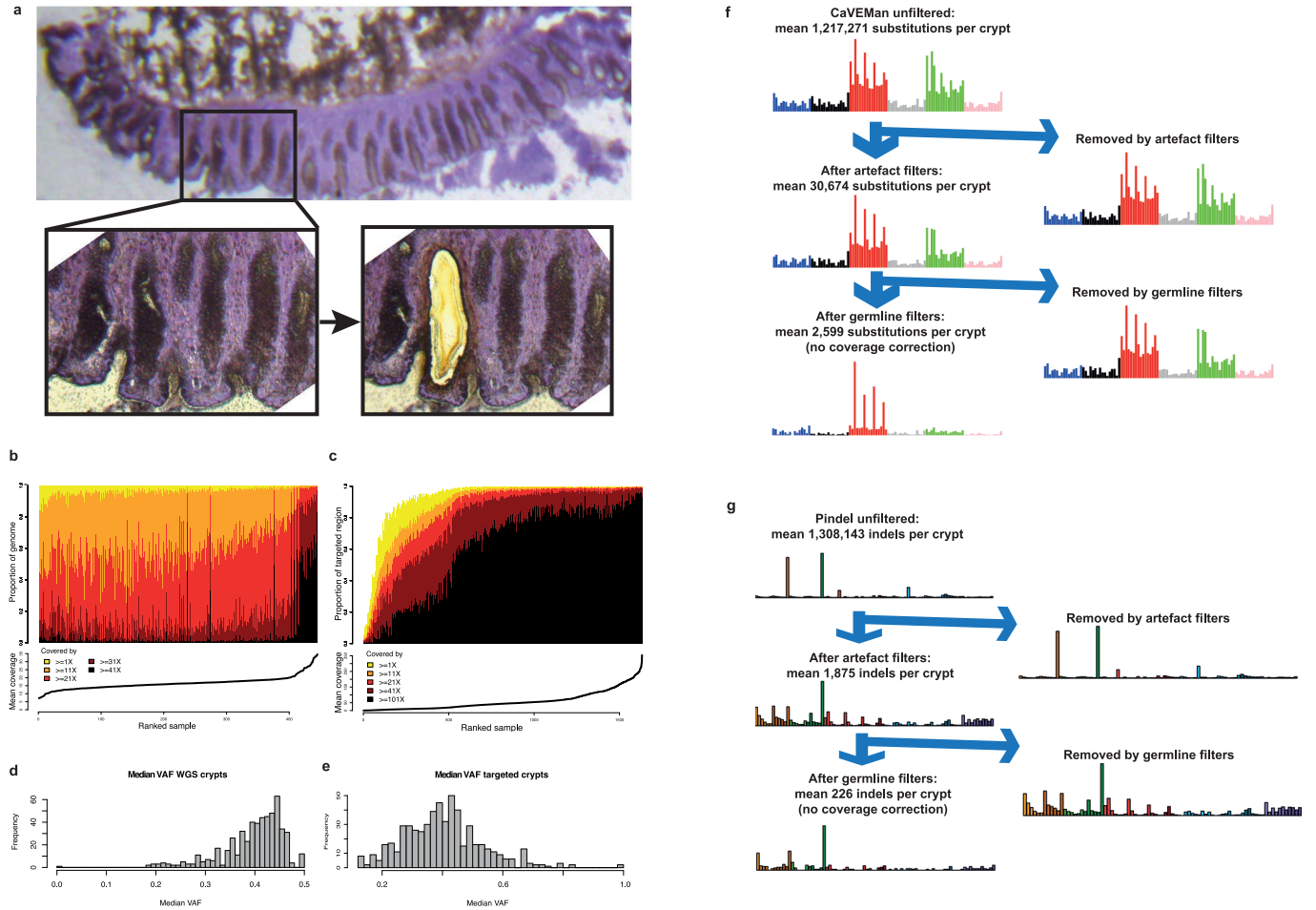


**Acknowledgements** This work was supported by the Wellcome Trust. We thank P. Scott, J. Fowler, D. Fernandez-Antoran and Y. Hooks for their advice with histology and laser-capture microdissection; M. Gerstung for his advice on statistics; the Sanger Institute Research and Development Facility for their help with sequencing microbiopsies; the staff of WTSI Sample Logistics, Genotyping, Pulldown, Sequencing and Informatics facilities for their contribution; K. Mahbubani, R. ten Hoopen, C. Scarpini and the Phoenix study team of N. Grehan, I. Debiram-Beecham, J. Crawte, T. Nukcheddy Grant, P. Lao-Sirieix and A. Hindmarsh for their help with sample collection; the Human Research Tissue Bank, which is supported by the National Institute for Health Research (NIHR) Cambridge Biomedical Research Centre from Addenbrooke's Hospital; and all the individuals who contributed samples to this study. Access to samples of transplants from organ donors was provided by the Cambridge Biorepository for Translational Medicine. A.N. was funded through an MRC Clinical Research Fellowship; the autopsy cohort was funded through this, an MRC core grant (RG84369) and an NIHR Research Professorship (RG67258) to R.C.F., and additional infrastructure support was provided from the CRUK-funded Experimental Cancer Medicine Centre in Cambridge.

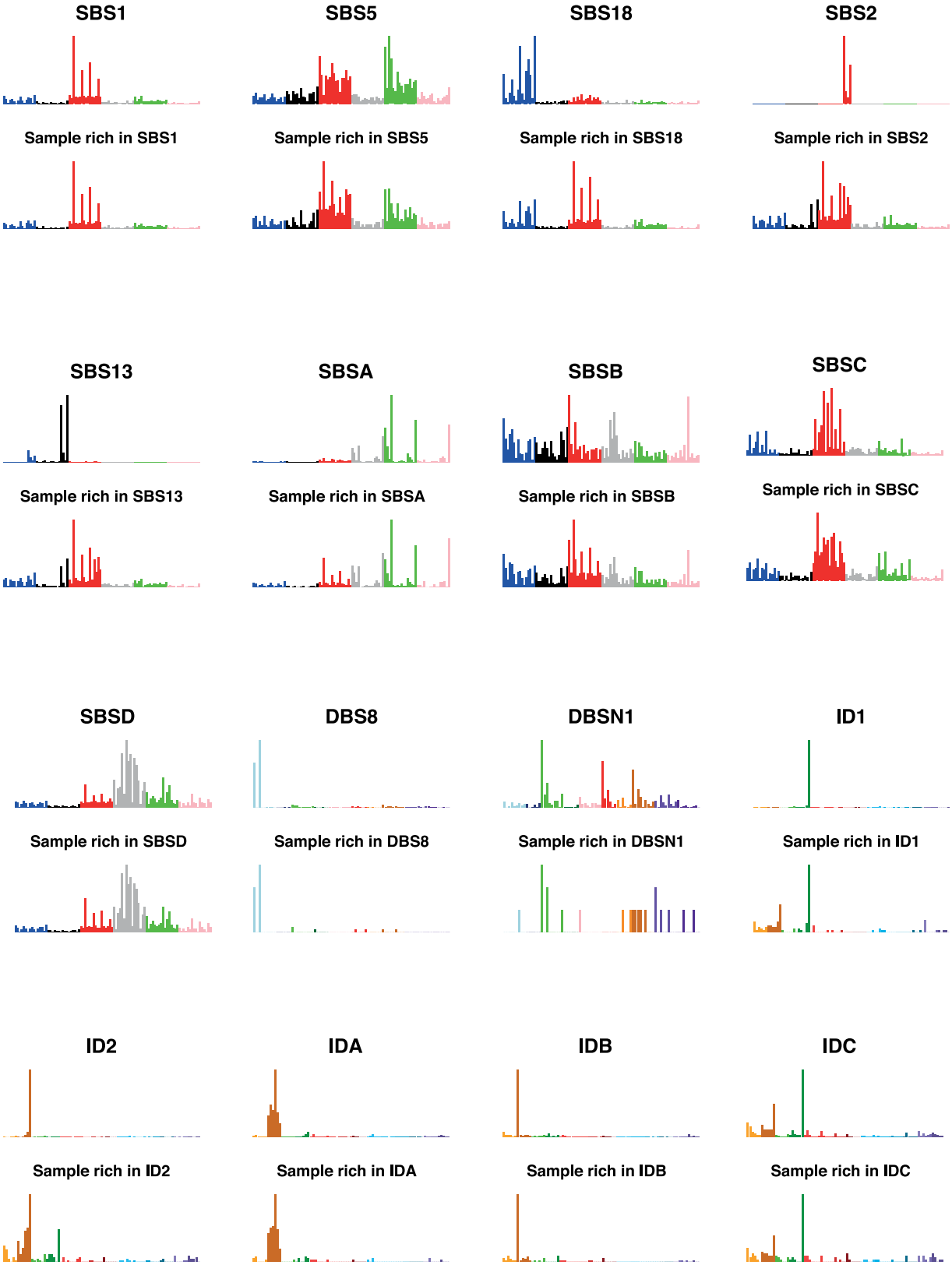

**Author contributions** M.R.S. and H.L.-S. designed the study and wrote the manuscript with contributions from all the authors; K.S.-P., N.C., M.Z., R.C.F., N.G., F.T., A.N., M.G. and L.M. recruited patients and obtained samples; P.E., R.J.O., H.L.-S. and L.M. devised the protocol for laser-capture microdissection and sequencing of colonic crypts; H.L.-S. prepared sections, microdissected and lysed colonic crypts; P.R. contributed to laser-capture microdissection; P.E. and C.A. made libraries; H.L.-S. performed most of the data curation and statistical analysis; S.O. estimated the rate of crypt fission; M.A.S. devised filters for substitution calling; J.W. performed in-house NMF signature extraction; T.H.H.C. and P.R. contributed to statistical analyses; L.O.N. provided technical assistance; P.J.C. and I.M. oversaw statistical analyses; and M.R.S. supervised the study.

**Competing interests** The authors declare no competing interests.

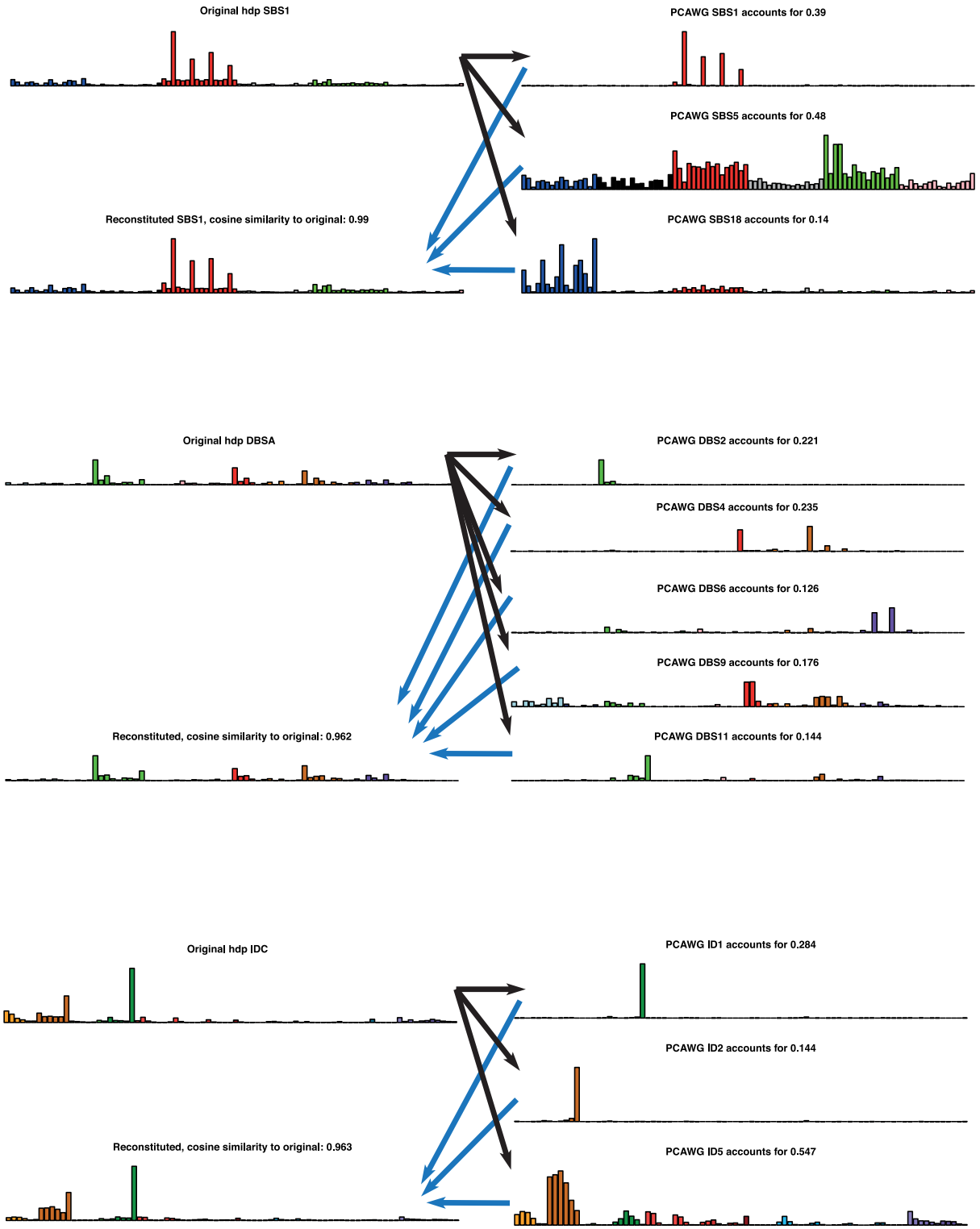

**Additional information**
**Supplementary information** is available for this paper at https://doi.org/10.1038/s41586-019-1672-7.
**Correspondence and requests for materials** should be addressed to M.R.S.
**Peer review information** *Nature* thanks Jacco van Rheenen and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.
**Reprints and permissions information** is available at http://www.nature.com/reprints.



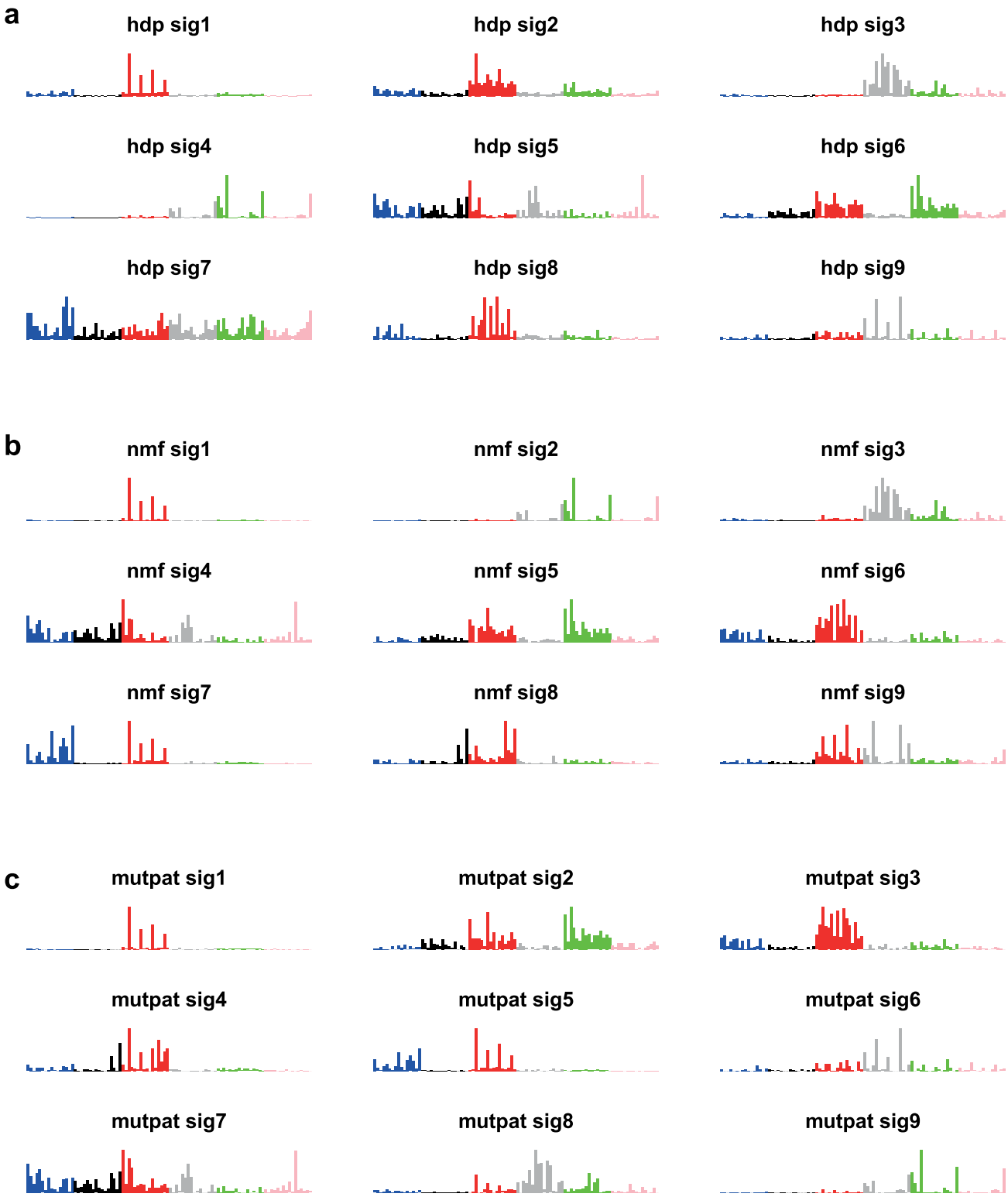

**Extended Data Fig. 1 | Laser-capture microdissection of crypts.** **a**, Representative image of a section of colonic tissue. The magnified inset shows the section before and after dissection of a crypt. **b**, **c**, Coverage of crypts that underwent whole-genome (**b**) and targeted (**c**) sequencing. **d**, **e**, VAFs (that is, is half of the clonal fraction) for crypts that underwent whole-genome (**d**) and targeted (**e**) sequencing. **f**, **g**, Substitutions (**f**) and indels (**g**) that were removed by filtering steps and their mutational spectra, arranged as in Fig. 1.



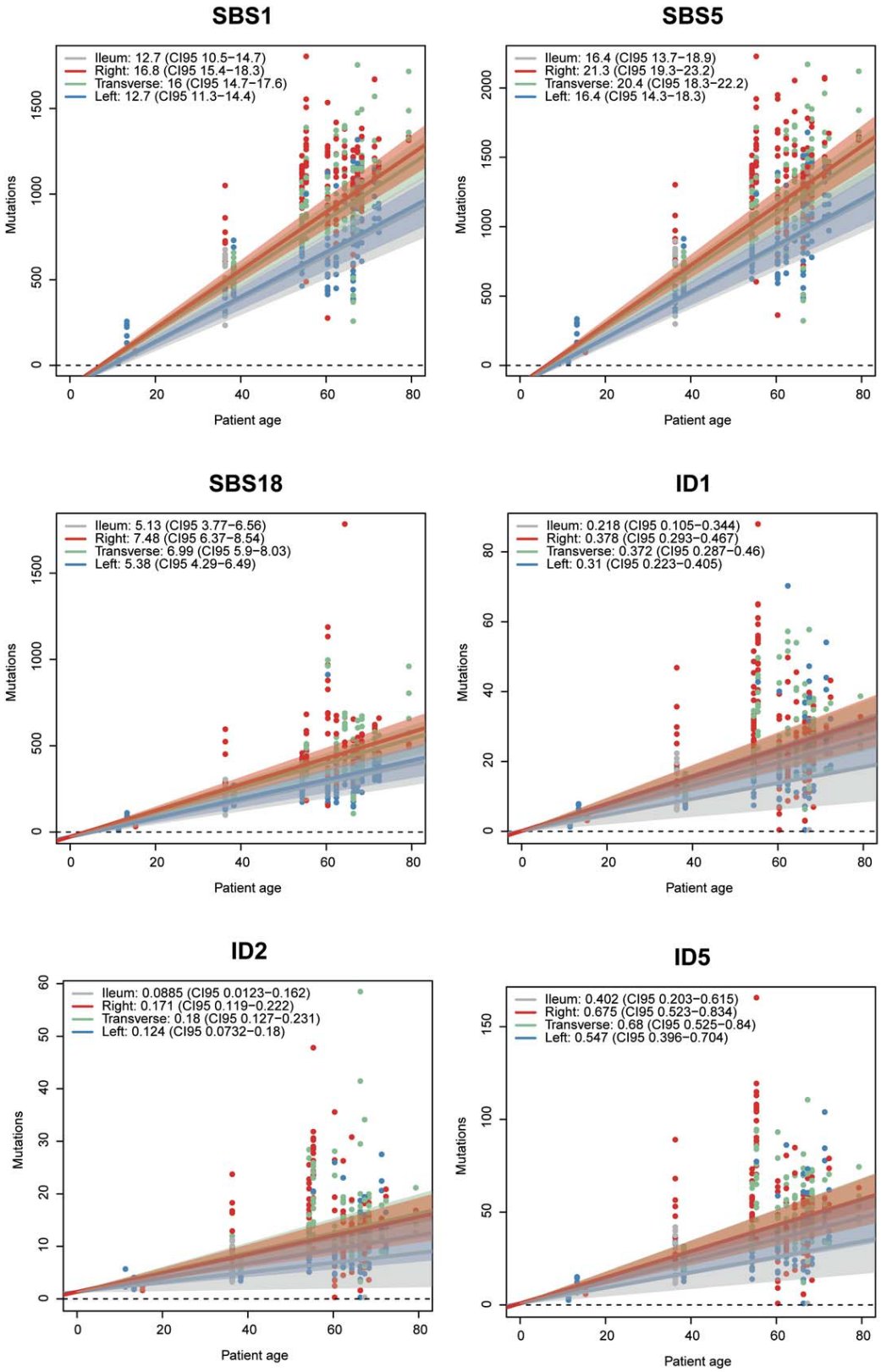

**Extended Data Fig. 2 | Results of HDP-based extraction of signatures.** Results of signature extraction using an HDP with pre-conditioning on signatures that are known to be active in colorectal cancer. For each signature, the extracted signature and the profile of a sample to which that signature contributes strongly are shown. Signatures are presented as in Fig. 2. The extraction of signatures using an HDP was followed by deconvolution by expectation maximization (Methods, Extended Data Fig. 3) to produce the versions of signatures that are presented in the main text.



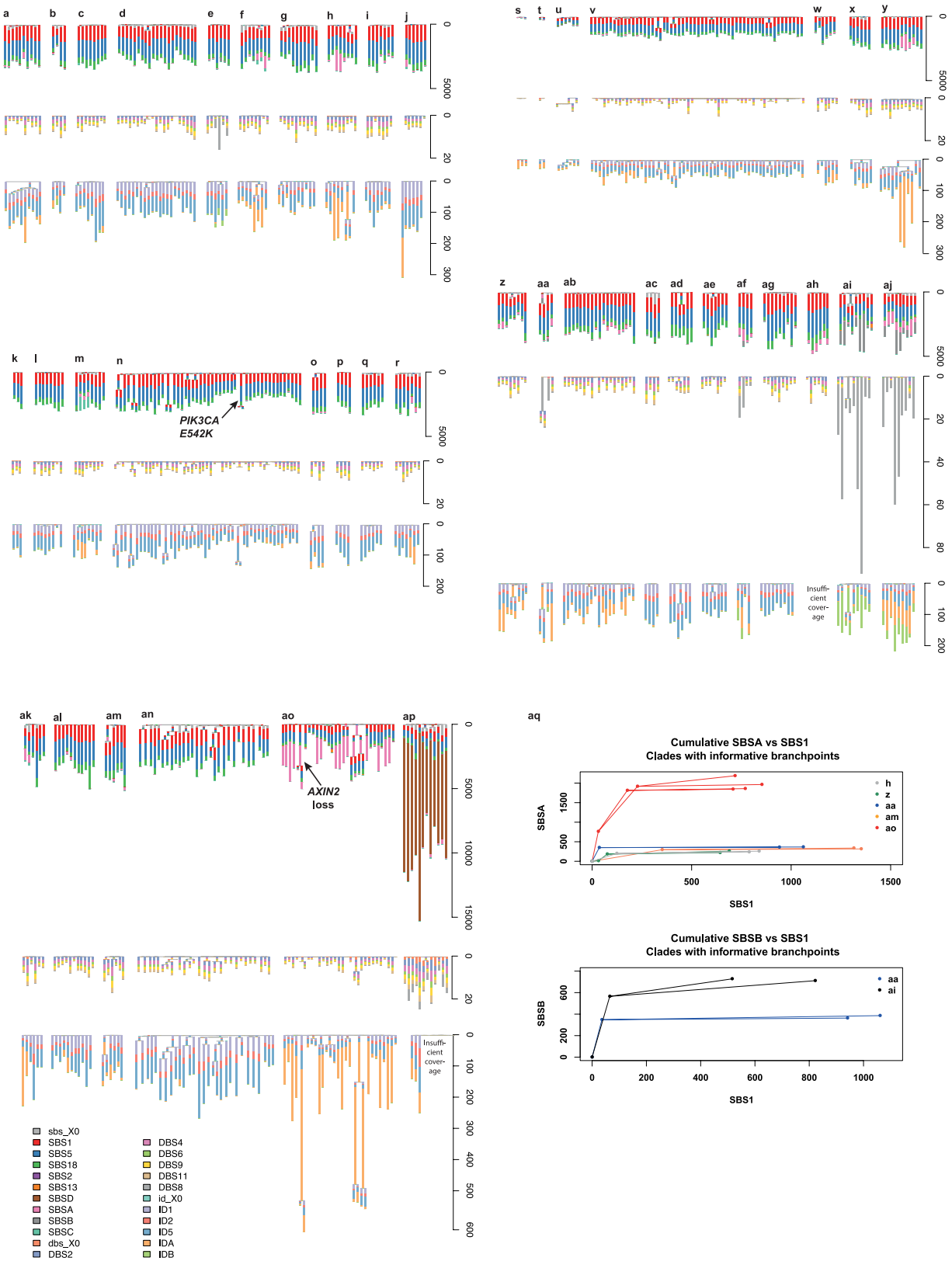

**Extended Data Fig. 3 | Decomposition of HDP signatures by expectation maximization.** Three signatures were decomposed (SBS1, DBSA and IDC). For each example, the original HDP version is shown on the top left, the PCAWG signatures that are deemed to contribute at least 10% of mutations to it on the right and the reconstituted signature that was built by combining the PCAWG signatures on the bottom left. The cosine similarity of the reconstituted signature to the original is shown.



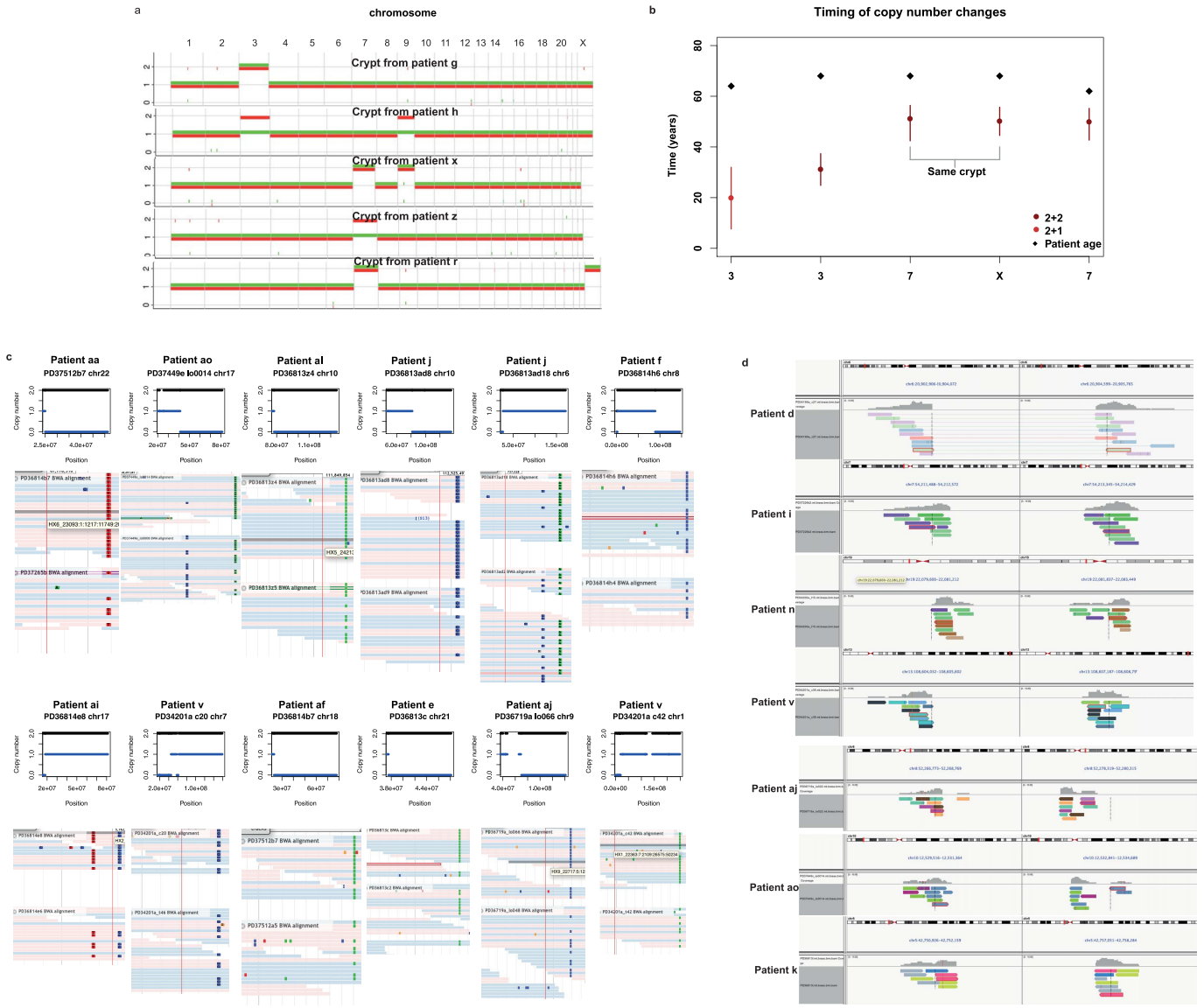

**Extended Data Fig. 4 | Validation of SBS signatures. a**–**c**, Other methods of signature extraction were run to test the robustness of signature decomposition. **a**, HDP without pre-conditioning on PCAWG. **b**, In-house NMF without pre-conditioning on PCAWG. **c**, NMF implemented by the MutationalPatterns package in R (Methods).



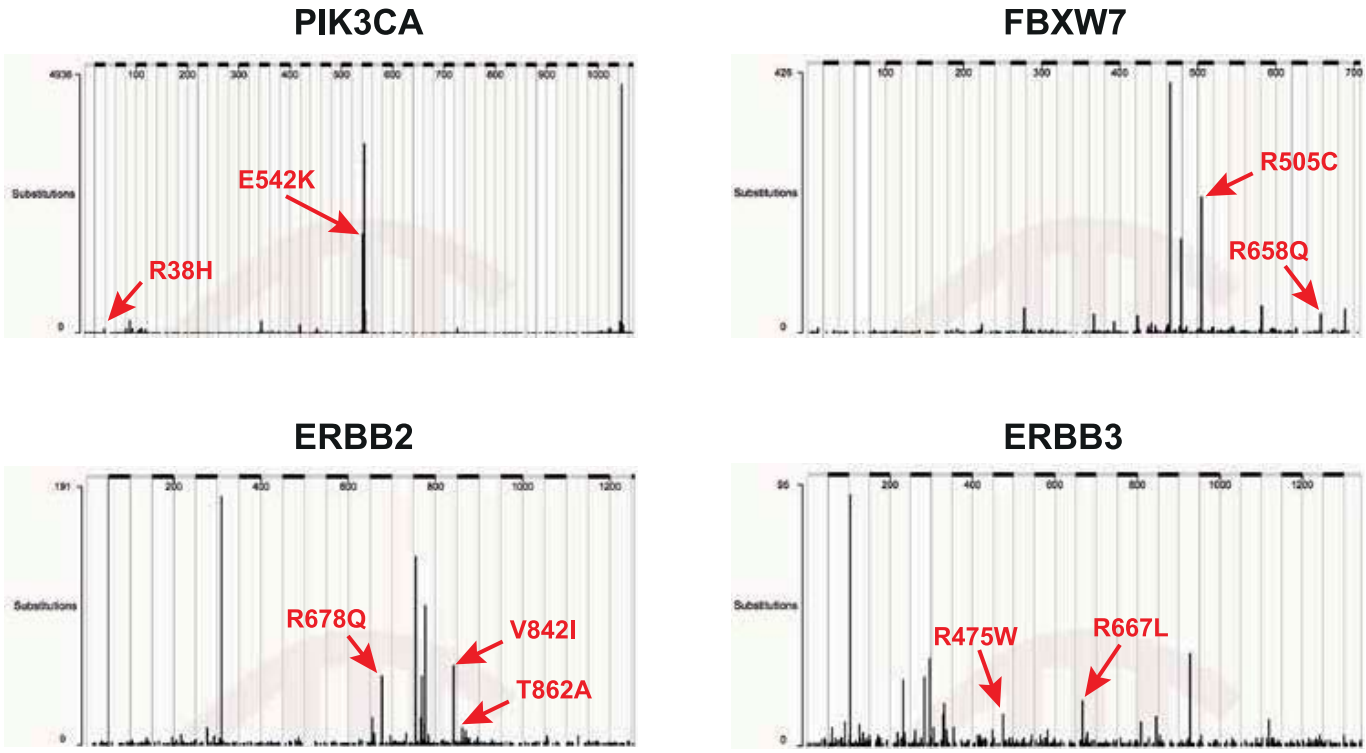

**Extended Data Fig. 5 | Linear modelling of the accumulation of signatures.** For signatures that appeared to show a linear accumulation with age, the mutation rate per site was determined using mixed models, in which age and site were used as fixed effects and individual as a random effect. Confidence intervals were determined by bootstrapping. $n = 445$ crypts from 42 individuals. Solid lines represent the mean slope of the regression and shaded areas its 95% confidence intervals (CI95).



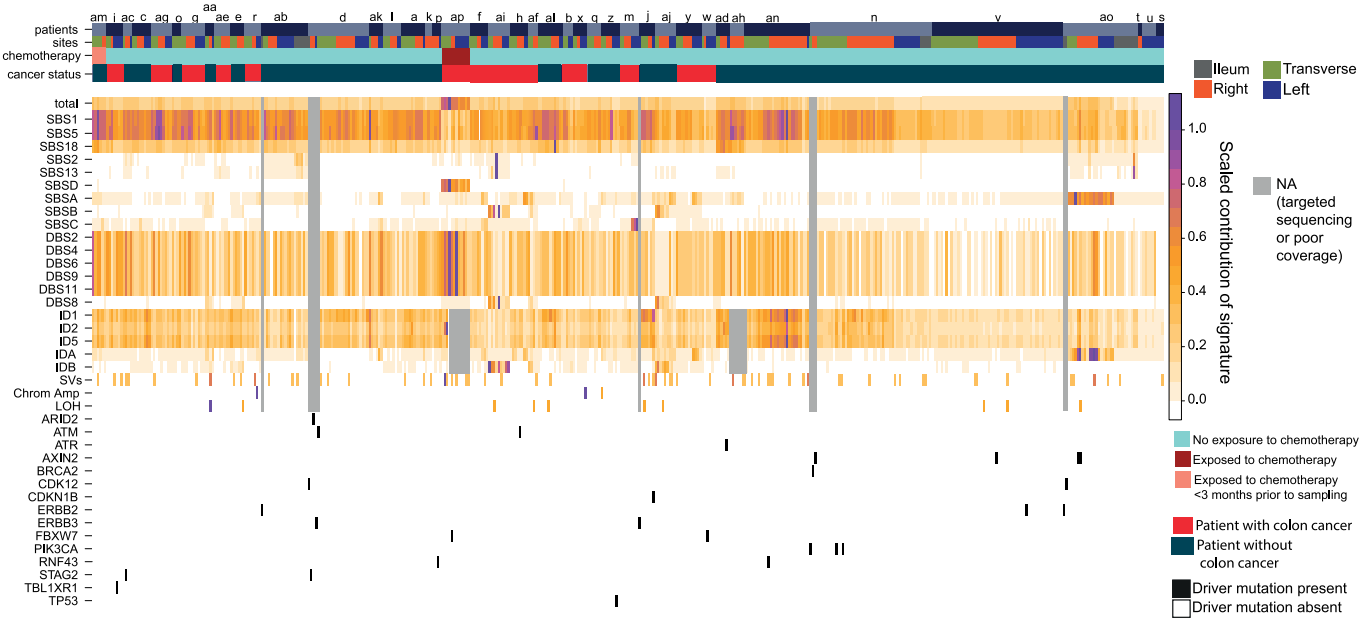

**Extended Data Fig. 6 | Crypt phylogenies. a–ap**, For each individual, the phylogeny of crypts is shown three times: at the top, with branch lengths proportional to the number of SBSs; in the middle, with branch lengths proportional to the number of DBSs; and on the bottom, with branch lengths proportional to the number of small indels. Scale bars are shown on the right. A stacked bar plot of the mutational signatures that contribute to each branch is overlaid over every branch. 'X0' indicates mutations that could not confidently be assigned to any signature. Note that the ordering of signatures along a given branch is just for visualization purposes; we cannot distinguish the timing of different signatures along a branch. **aq**, The cumulative burden of SBSA (top) and SBSB (bottom) is plotted relative to the cumulative burden of SBS1 to time these mutational processes throughout life. Informative clades are shown (from patients labelled as in the rest of the figure), with every node and tip of the clade plotted in the space of the cumulative number of mutations that are due to a given signature that have occurred up until that node in the tree. Lines represent the branching structure of the tree.



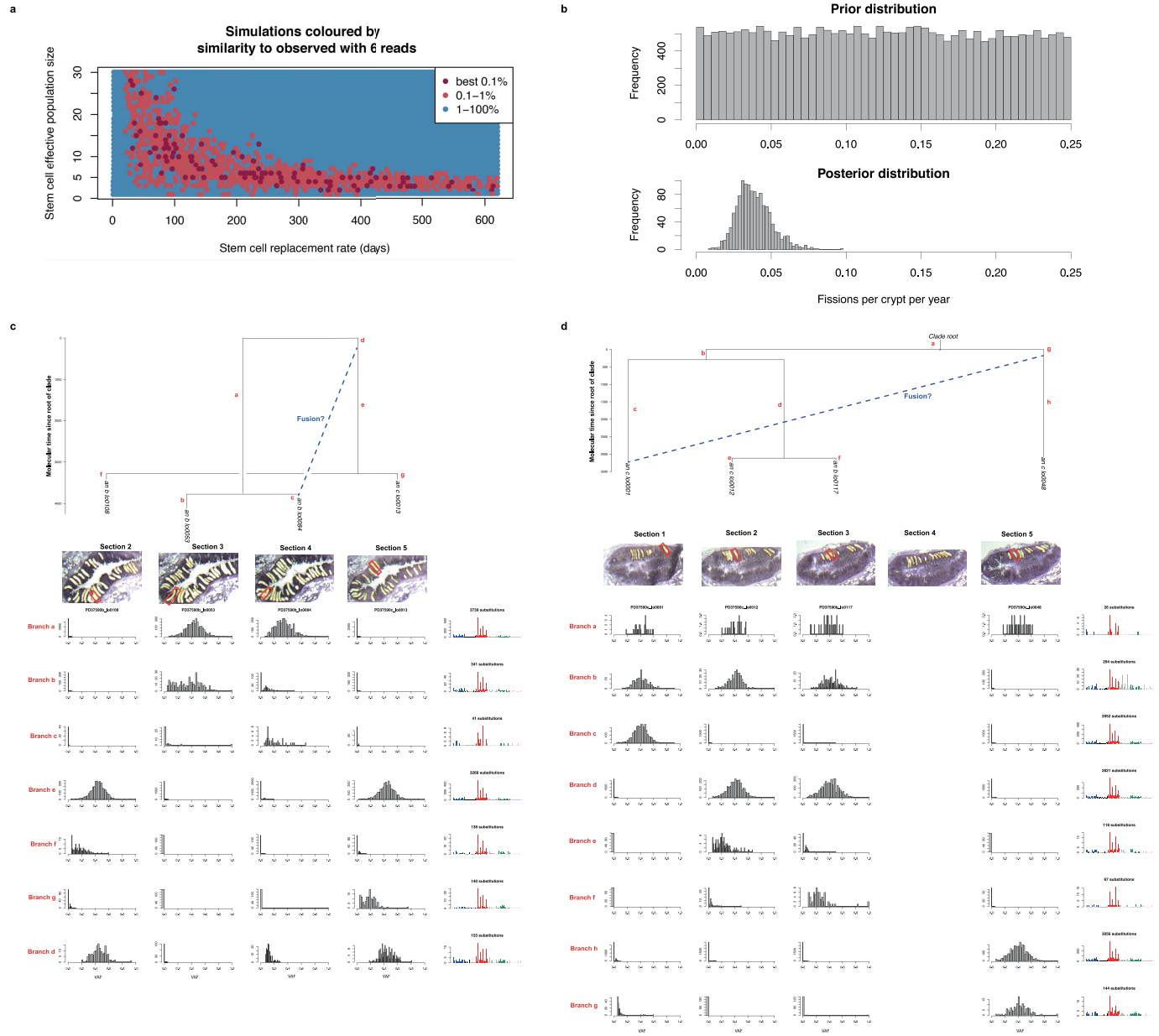

**Extended Data Fig. 7 | Copy-number changes and structural variants in normal colon. a**–**d**, A total of 449 crypts had sufficient coverage to be evaluated. **a**, Whole-chromosome amplifications in five crypts. The copy-number state (*y* axis) for each chromosome is shown, with one allele coloured red and the other green. Chromosomes are labelled along the top of the graph. **b**, Timing of copy-number changes throughout life. Vertical bars represent 95% confidence intervals, which were determined by bootstrapping. Horizontal bars represent the most likely time of the copy-number change, as defined by mutationTimeR (see Supplementary Information). **c**, Crypts with loss of heterozygosity (LOH). For each chromosome with a LOH event, the copy number across the whole chromosome is shown at the top, with the total copy number in black and the copy number for the minor allele in blue. The images at the bottom show example SNPs that support the LOH. In each case, reads from the crypt in question are shown above, and reads from its matched normal below. Thus, in the first image, the wild-type state (below) is heterozygous for a T SNP (red), whereas in the crypt in question (above), this polymorphism has now become homozygous. Small deviations from a fully homozygous state are probably a result of stromal contamination. **d**, Reads supporting structural variants in normal colon. Patients are labelled as in Extended Data Fig. 6.



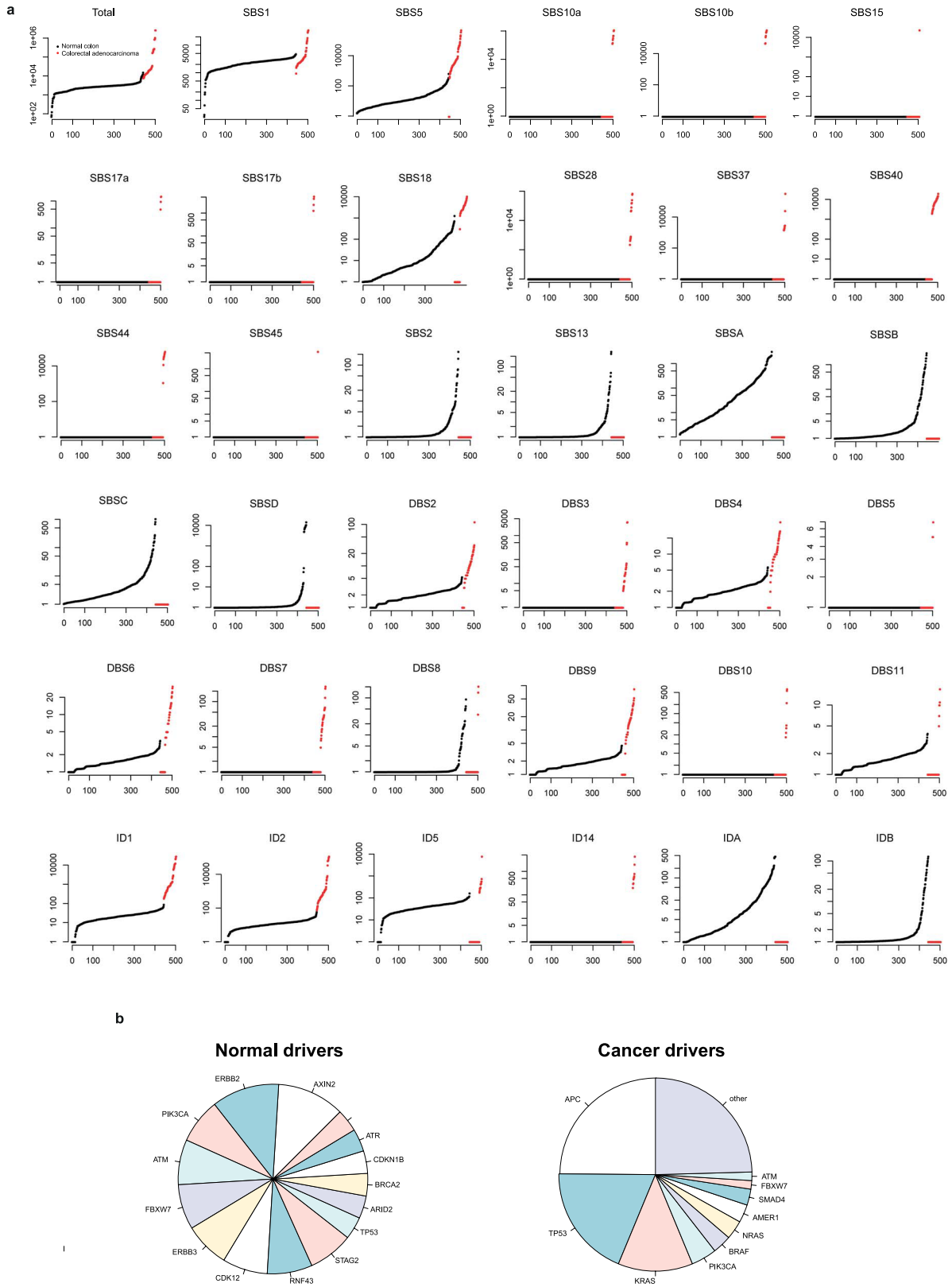

**Extended Data Fig. 8 | Gain-of-function driver mutations in normal colon.** Putative driver missense mutations in oncogene hotspots. The number of substitutions catalogued in COSMIC[53] is shown on the *y* axis at each position along the gene, with the mutations that were observed in our cohort indicated with arrows.




**Extended Data Fig. 9 | Occurrence matrix of signatures and driver mutations in crypts.** For all crypts that were whole-genome sequenced to sufficient depth, and for crypts that underwent targeted sequencing and in which driver mutations were found, the signatures and driver mutations are shown. Each vertical column represents a crypt. The individual to whom each crypt belongs is indicated by the alternating colours in the top bar (labelling as in Extended Data Fig. 6). The site to which each crypt belongs is shown underneath. The matrix is coloured by the contribution of each signature to each crypt, normalized for each signature: the crypt with the largest contribution of a given signature is purple and the crypt with the smallest contribution is white. Crypts in which the signatures could not be assessed, either because they underwent targeted sequencing or because the coverage was poor, are grey. Driver mutations, including heterozygous mutations in tumour suppressor genes, are indicated by a black bar.




**Extended Data Fig. 10 | Stem cell dynamics of normal colon. a**, Number of stem cells and replacement rate of stem cells in normal human colonic crypts, as estimated by approximate Bayesian computation. Each point represents a simulation. Points are coloured according to their similarity to the observed data: the most similar 0.1% are coloured dark red, and so on, until the least similar simulations are blue. **b**, Approximate Bayesian computation of the rate of crypt fission (fissions per crypt per year) in the human colon. The prior distribution of the crypt fission rate (which was used to simulate many biopsies of the colon) is shown above, and the posterior distribution of the crypt fission rate (estimated by neural network regression on the simulations) is shown below. **c**, **d**, Evidence of crypt fusion in human colon. In each case, a phylogeny is shown at the top that depicts the genetic relationships between selected crypts. Dashed blue lines show mutations with a low allele fraction that are shared between crypts in a manner incompatible with the phylogeny dictated by the clonal mutations. Below each crypt in the phylogeny is an image that depicts its position in the section. Sections are labelled according to their *z*-stacked order. The allele fractions of mutations on each branch of the phylogeny in each crypt are shown at the bottom. The trinucleotide context of the mutations that occurred on each branch is shown on the right. See also Supplementary Information.




**Extended Data Fig. 11 | Comparison of the mutational signatures and driver landscape of normal crypts and colorectal adenocarcinomas. a**, Comparison of the burden of mutations for every mutational signature. For each signature, the *y* axis shows the mutational burden + 1 of every sample on a logarithmic scale. Normal colon and cancer samples are ordered within their groups. The signature attributions and mutational burden for colorectal adenocarcinoma are from a previous study[2]. A total of 60 cancers are compared with 472 normal crypts. **b**, The proportion of driver mutations in each gene in normal colon (left) and colorectal cancer (right). The frequency of driver mutations in cancer was derived using data from TCGA research network[43] (Supplementary Methods).

nature research

Corresponding author(s): *Double-blind peer review submissions: write DBPR and your manuscript number here instead of author names.*

Last updated by author(s): *YYYY-MM-DD*

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☐ | ☒ | A description of all covariates tested |
| ☐ | ☒ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☒ | ☐ | For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☐ | ☒ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒ | ☐ | Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |

*Our web collection on statistics for biologists contains articles on many of the points above.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | No code was used to collect data in this study - code was only used to analyse data. |
| Data analysis | All code central to this study is available on github at https://github.com/HLee-Six/colon_microbiopsies, or is included in the supplementary results. Any additional code will be provided by the authors on request.<br><br>Additional publicly available code was used:<br>BWA-Mem - version 0.7.12-r1039<br>Caveman - version 1.11.2<br>Pindel - version 2.1.0<br>ASCAT - version 4.0.1<br>Brass - version 6.0.5<br>MutationTimeR - version 0.99.1<br>TelomereCat - version 3.1.2 |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Whole genome and targeted sequencing data are deposited in the European Genome Phenome Archive (EGA). sequencing data have been deposited with EGA accession EGAD00001004192 and EGAD00001004193.
There are no restriction on data availability.

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Sample size | Sample size was chosen based on the resources available. As this was a description of the landscape, rather than an attempt to test a particular hypothesis, a power calculation is not applicable. Our sample size was an order of magnitude larger than previous studies, which we hoped would be sufficient to yield new insights. |
| Data exclusions | Some crypts were excluded if their sequencing coverage and/or clonality were too poor for mutations to be called accurately.<br>10X coverage was required for calling copy number changes.<br>For statistical analyses in supplementary results 2, only crypts whose median depth multiplied by their median variant allele fraction was greater than 3 were included.<br>These specific criteria were not pre-established, although it had always been expected than some coverage and clonality cut-offs would be required to obtain good quality data. |
| Replication | Our article describes the genomic landscape of the normal colon, it does not test specific hypotheses, and so replication does not apply in its usual way.<br>Sequencing replicates are not normally possible as once a crypt has been sequenced it has been used up. |
| Randomization | Our article describes the genomic landscape of the normal colon, it does not test a treatment, and so randomisation does not apply. |
| Blinding | Our article describes the genomic landscape of the normal colon, it does not test a treatment, and so blinding does not apply. |

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ Antibodies |
| ☒ | ☐ Eukaryotic cell lines |
| ☒ | ☐ Palaeontology |
| ☒ | ☐ Animals and other organisms |
| ☐ | ☒ Human research participants |
| ☒ | ☐ Clinical data |

### Methods

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ ChIP-seq |
| ☒ | ☐ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |

## Human research participants

Policy information about studies involving human research participants

| | |
|---|---|
| Population characteristics | We obtained healthy colonic biopsies from four cohorts (Supplementary Table 1). The first represents seven deceased organ donors ranging in age from 36 to 67, from whom colonic and small intestinal biopsies were taken at the time of organ donation |

(REC 15/EE/0152). The second represents individuals aged 60 to 72 who were having a colonoscopy following a positive faecal occult blood test as part of the Bowel Cancer Screening Programme (Ethical approval 08-H0308-13); we selected 16 who were not found to have either an adenoma or a carcinoma on colonoscopy, and 15 who were found to have a colorectal carcinoma (the normal biopsies that we use were distant from these lesions). The third cohort represents three paediatric patients who underwent routine colonoscopy to exclude inflammatory bowel disease and who were found to have a completely normal intestinal mucosa macroscopically and histologically (REC 12/EE/0482). The final cohort included one 78 year-old gentleman with oesophageal cancer who underwent a warm autopsy (REC 13/EE/0043). This gentleman had been treated with palliative chemotherapy of Epirubicin, Oxaliplatin and Capecitabine within the three months before the autopsy; given the slow monoclonal conversion within crypts, and mutations due to these chemotherapies are unlikely to be detected. All samples were obtained with informed consent and studies approved by East of England Research Ethics Committees.

Recruitment

The transplant donor cohort recruits those who are on the organ donor register and are suitable for transplantation. Donors must therefore have healthy organs. This should not affect the results of a survey or normal colonic epithelium.

In the Bowel Cancer Screening programme, men and women aged 60-72 were offered a faecal occult blood test every 2 years. It is possible that there may be higher uptake by those who are more concerned about their health and therefore slightly healthier than the general population. Those who had a positive FOB are more likely to have colonic pathology, but all were found to have no cancer near the lesion at colonoscopy. On the whole, the recruitment of this cohort should not affect our results.

One warm autopsy patient with oesophageal cancer consented for his body to be used for research as part of the Phoenix study. Without this sort of recruitment it would be very difficult to capture the extremes of age in our study.

Children who had gastrointestinal symptoms suggestive of inflammatory bowel disease but a normal colonoscopy were included. While, as they had symptoms, they may be slightly less likely to have normal colons than the rest of the population, the colonoscopy excludes anything obvious. The most likely cause for their symptoms is infective, which is unlikely to affect the genomic landscape of the colon.

Ethics oversight

East of England Research Ethics community

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Article

# Somatic mutations and clonal dynamics in healthy and cirrhotic human liver

https://doi.org/10.1038/s41586-019-1670-9

Received: 17 November 2018

Accepted: 12 September 2019

Published online: 23 October 2019

Simon F. Brunner[1], Nicola D. Roberts[1], Luke A. Wylie[1], Luiza Moore[1], Sarah J. Aitken[2,3], Susan E. Davies[3], Mathijs A. Sanders[1,4], Pete Ellis[1], Chris Alder[1], Yvette Hooks[1], Federico Abascal[1], Michael R. Stratton[1], Inigo Martincorena[1], Matthew Hoare[2,5]* & Peter J. Campbell[1,6]*

The most common causes of chronic liver disease are excess alcohol intake, viral hepatitis and non-alcoholic fatty liver disease, with the clinical spectrum ranging in severity from hepatic inflammation to cirrhosis, liver failure or hepatocellular carcinoma (HCC). The genome of HCC exhibits diverse mutational signatures, resulting in recurrent mutations across more than 30 cancer genes[1–7]. Stem cells from normal livers have a low mutational burden and limited diversity of signatures[8], which suggests that the complexity of HCC arises during the progression to chronic liver disease and subsequent malignant transformation. Here, by sequencing whole genomes of 482 microdissections of 100–500 hepatocytes from 5 normal and 9 cirrhotic livers, we show that cirrhotic liver has a higher mutational burden than normal liver. Although rare in normal hepatocytes, structural variants, including chromothripsis, were prominent in cirrhosis. Driver mutations, such as point mutations and structural variants, affected 1–5% of clones. Clonal expansions of millimetres in diameter occurred in cirrhosis, with clones sequestered by the bands of fibrosis that surround regenerative nodules. Some mutational signatures were universal and equally active in both non-malignant hepatocytes and HCCs; some were substantially more active in HCCs than chronic liver disease; and others—arising from exogenous exposures—were present in a subset of patients. The activity of exogenous signatures between adjacent cirrhotic nodules varied by up to tenfold within each patient, as a result of clone-specific and microenvironmental forces. Synchronous HCCs exhibited the same mutational signatures as background cirrhotic liver, but with higher burden. Somatic mutations chronicle the exposures, toxicity, regeneration and clonal structure of liver tissue as it progresses from health to disease.

Identifying somatic mutations in non-malignant tissue requires approaches that overcome the polyclonality of this tissue, such as single-cell sequencing[9], cultures of single cells[8,10] or microbiopsy sequencing[11]. The latter relies on local cell division with limited migration leading to a clonal patchwork, which has been observed in liver tissue[12]. We generated whole-genome sequences from 482 laser-capture microdissections (LCMs) of 100–500 hepatocytes (Extended Data Fig. 1a) across 14 individuals: 5 healthy controls; 4 patients with cirrhosis from alcohol-related liver disease (ARLD) and 5 patients with cirrhosis from non-alcoholic fatty liver disease (NAFLD) (Supplementary Tables 1, 2, Extended Data Figs. 4–6). Samples of normal liver were acquired from hepatic resections of colorectal cancer metastases, and samples of cirrhotic liver were taken from patients who underwent liver transplants for synchronous but distant HCC.

To evaluate sensitivity and specificity, we generated independent libraries and sequencing data from different sections of the same biopsy, microdissecting the same $x$, $y$-region from adjacent $z$-stacks separated by around 20 μm. Concordance was high between variants that were called in adjacent sections, but not between distant pairs, suggesting that the specificity of mutation calls was high (Extended Data Fig. 1b). Sensitivity across patients ranged from 50 to 95%, depending on coverage and clonality (Extended Data Fig. 1c–f). As a further check on specificity, targeted deep sequencing of cancer genes from the same library as 96 whole-genome samples confirmed 16 of the 17 mutations that were originally called. In keeping with polyploidy as a late stage of differentiation in the liver[13], 20–25% of mature hepatocytes in microdissected samples were multinuclear (Extended Data Fig. 1g). We therefore deployed copy-number algorithms with an expected ploidy of 4, and report mutational burdens per diploid genome, rather than per cell.

We observed considerable heterogeneity in the burden of somatic substitutions both between and within patients (Fig. 1a, Supplementary Tables 3, 4). Using mixed-effects models, microdissections from

[1]Cancer Genome Project, Wellcome Trust Sanger Institute, Hinxton, UK. [2]CRUK Cambridge Institute, Cambridge, UK. [3]Department of Pathology, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK. [4]Department of Hematology, Erasmus University Medical Center, Rotterdam, The Netherlands. [5]Department of Medicine, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK. [6]Department of Haematology and Stem Cell Institute, University of Cambridge, Cambridge, UK. *e-mail: Matthew.Hoare@cruk.cam.ac.uk; pc8@sanger.ac.uk

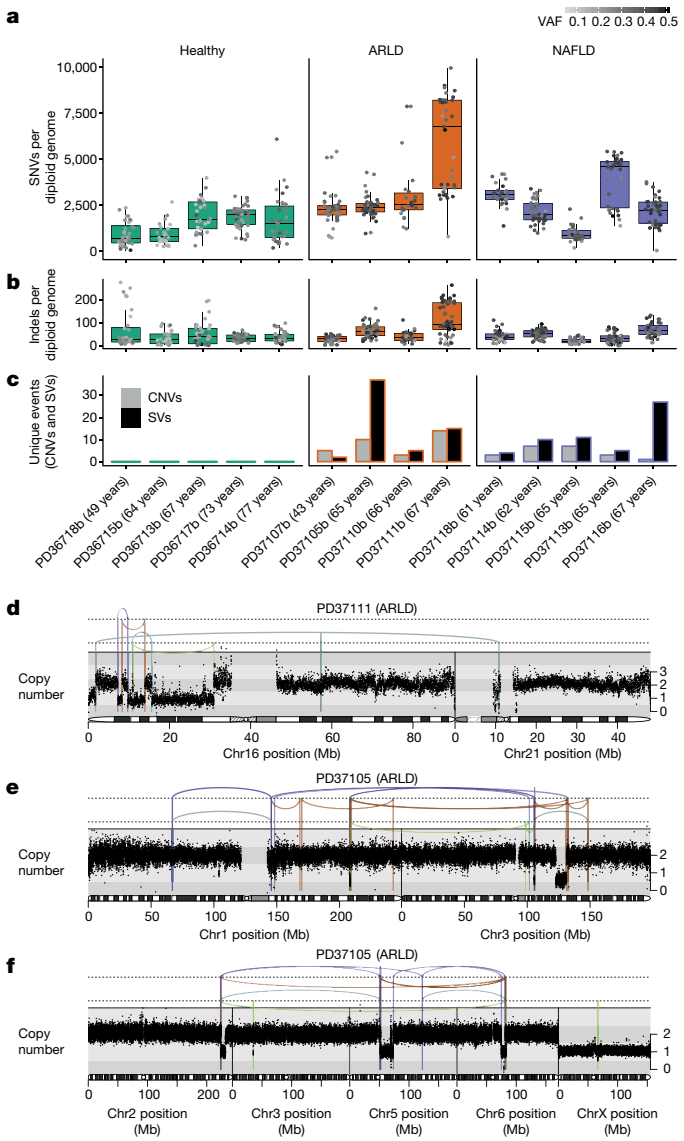



**Fig. 1 | Mutational burden observed in non-cancerous hepatocytes.** **a**, Burden of single-nucleotide variants (SNVs), corrected by sensitivity of mutation detection. Each box plot represents a patient ($n = 14$ patients; 482 microdissections) and each dot represents one laser-capture microdissected sample. The grey-to-black intensity of the points reflects the median variant allele fraction (VAF) of mutations in each microdissection. Boxes in the box plots indicate median and interquartile range; whiskers denote range. **b**, Burden of indel variants ($n = 14$ patients; 482 microdissections). **c**, Burden of copy-number variants (CNVs) and structural variants (SVs), represented as the number of unique events per patient. **d**, Chromothripsis involving chromosomes 16 and 21, observed in patient PD37111. Black points represent corrected read depth along the chromosome. Lines and arcs represent structural variants, coloured by the orientation of the joined ends (purple, tail-to-tail inverted; brown, head-to-head inverted; turquoise, tandem-duplication-type orientation; green, deletion-type orientation). **e**, Chromothripsis involving chromosomes 1 and 3, observed in patient PD37105. **f**, Chromothripsis involving chromosomes 2, 5 and 6, observed in patient PD37105 (in a separate clone to **e**).

cirrhotic livers had, on average, 1,251 (95% confidence interval, 233–2,268; $P = 0.02$) extra substitutions per diploid genome compared to normal livers, independent of age. In accordance with published values[8], the estimated rate of accumulation of mutations was 33 per year per diploid genome, albeit with wide confidence intervals (95% confidence interval, −17 to 84; $P = 0.18$) and moderate variation between individuals (estimated between-individual s.d., 13 per year). Insertions and deletions (indels) showed the same heterogeneity between and within individuals as substitutions (Fig. 1b).

Structural variants and copy-number alterations occurred in moderate numbers across all nine patients with liver cirrhosis, despite being rare in normal liver (Fig. 1c, Extended Data Fig. 2, Supplementary Tables 3, 4). Occasional aneuploidy at whole-chromosome or arm level occurred, as well as focal events including deletions, tandem duplications and unbalanced translocations (Extended Data Fig. 2). We found five separate clusters of structural variants across three patients, with patterns indicative of chromothripsis[14] (Fig. 1d–f, Extended Data Fig. 2). Chromothripsis—in which multiple rearrangements occur in a single catastrophic mitosis[14]—is a major process of mutation in cancers (occurring in around 5% of HCCs[15]), but is rare in normal somatic cells. Our observation of 1–2% of clones with chromothripsis in chronic liver disease suggests that sustained toxicity and regeneration substantially increases mitotic stress in hepatocytes.

We screened for driver mutations among coding regions, 5′-untranslated regions (UTRs), 3′-UTRs and promoters (Supplementary Tables 5–8). No elements were significant genome-wide after correcting for multiple hypotheses, so we focused on the 30 most-prevalent HCC genes[1–5]. These carried 22 non-synonymous variants that were seen in both normal and cirrhotic samples and included inactivating mutations in the tumour suppressor genes *ACVR2A*, *ARID2*, *ARID1A* and *TSC2* (Extended Data Fig. 3a). When hypothesis testing was restricted to these 30 genes, *ALB* and *ACVR2A* were significant ($q = 0.001$ and $q = 0.001$, respectively). Recurrence in *ALB* (which encodes the protein albumin) probably reflects a mutational process in which indels preferentially occur in highly expressed genes, as reported in HCCs[5,16] (Extended Data Fig. 3b, c). Assuming no negative selection, we can use the ratio of non-synonymous to synonymous substitutions for the 30 HCC genes to estimate the number of driver substitutions among them[17]; this gives a 95% confidence interval of 0.0–13.2 driver mutations in total across 482 microdissections (that is, less than 3%). Among copy-number aberrations of potential importance[1,2,18] (Supplementary Table 9), we found instances of loss of chromosomes 22 and 8p, and gain of chromosome 8q. Two focal deletions in different patients spanned *ACVR2A* (Extended Data Fig. 2c, e). We also found a reciprocal inversion that deleted *CDKN2A* (Extended Data Fig. 2f), the most common focal deletion in HCC, and a deletion that affected *ARID5A*.

We reconstructed phylogenetic trees[19] and layered them onto the histology of the specimens. Samples from healthy controls showed the highly polyclonal nature of normal liver, with little genetic relatedness among even closely located microdissections (Fig. 2a–d, Extended Data Fig. 4). Samples from patients with chronic liver disease showed a clonal structure that was more complex, from which three general inferences can be drawn (Fig. 2e–p, Extended Data Figs. 5, 6). First, we found no sharing of mutations between adjacent liver nodules separated by fibrotic bands. This suggests that the connective tissue that is laid down during cycles of damage and regeneration sequesters clones from the early stages of the disease process. Second, some cirrhotic nodules were monoclonally derived (Fig. 2j, n, for example), whereas others were oligoclonal (Fig. 2f), and shared mutations often extended across microdissections that were millimetres apart. Third, branching structures in phylogenies point to subclonal diversification within nodules. Within such a clone, the proportion of shared, clonal mutations on the trunk relative to those on the subclonal branches gives an estimate in molecular time of when the most-recent common ancestor of the clone emerged. In some patients (for example, patient PD37114; Fig. 2i, j), the common ancestor of individual nodules emerged relatively early in molecular time, whereas in others (for example, patient PD37116; Fig. 2m, n), the common ancestor appeared much more recently. As the majority of liver cells do not have driver mutations, the size and rapidity of clonal expansions observed here demonstrate the considerable intrinsic capacity of hepatocytes to regenerate in response to liver damage.

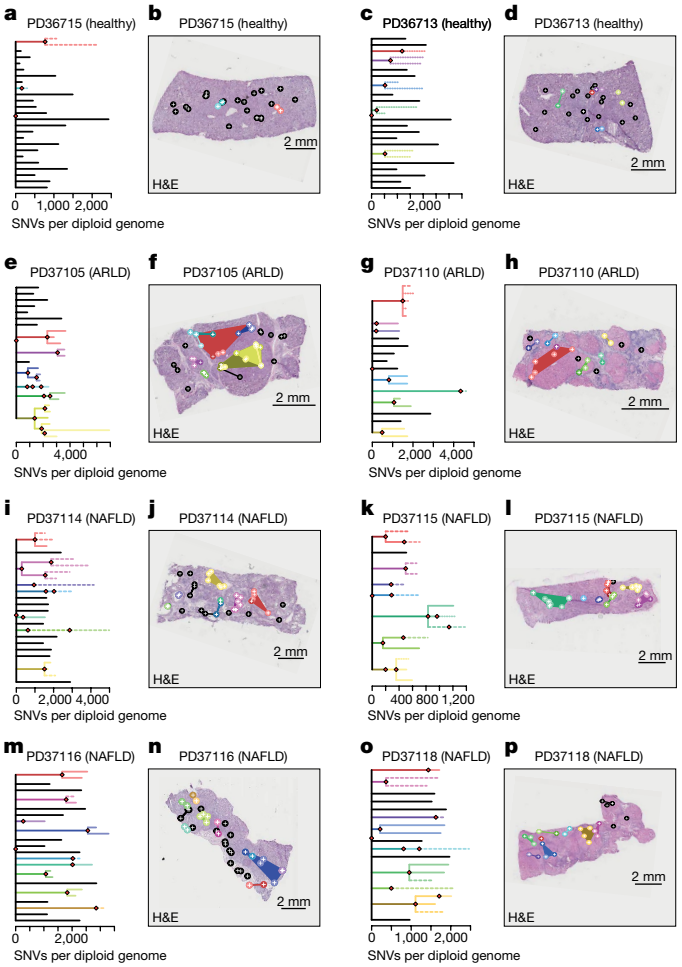



**Fig. 2 | Phylogenetic reconstruction of hepatocyte clones. a**, Phylogenetic tree constructed from clustering of mutations across microdissected samples in a healthy individual (PD36715). Lengths of branches (*x* axis) indicate the numbers of mutations assigned to that branch. Solid lines indicate that nesting is in accordance with the pigeonhole principle; dashed lines indicate that nesting is in accordance with the pigeonhole principle, assuming that hepatocytes represent 70% of cells; dotted lines indicate that nesting is only based on clustering (clones are assigned as nested if the VAFs of constituent microdissections are lower than those in the parental clone). **b**, Representation of branches from the phylogenetic tree in **a** according to their physical coordinates, overlaid onto a haematoxylin and eosin (H&E)-stained section. Black points represent branches of the tree that share no mutations with any other samples; coloured points represent branches with shared clonal relationships ($n = 26$ microdissections). **c**, **d**, A second healthy individual (PD36713; $n = 30$ microdissections). **e**, **f**, Patient with ARLD (PD37105; $n = 31$ microdissections). **g**, **h**, Patient with ARLD (PD37110; $n = 22$ microdissections). **i**, **j**, Patient with NAFLD (PD37114; $n = 41$ microdissections). **k**, l, Patient with NAFLD (PD37115; $n = 34$ microdissections) **m**, **n**, Patient with NAFLD (PD37116; 43 microdissections). **o**, **p**, Patient with NAFLD (PD37118; 26 microdissections).

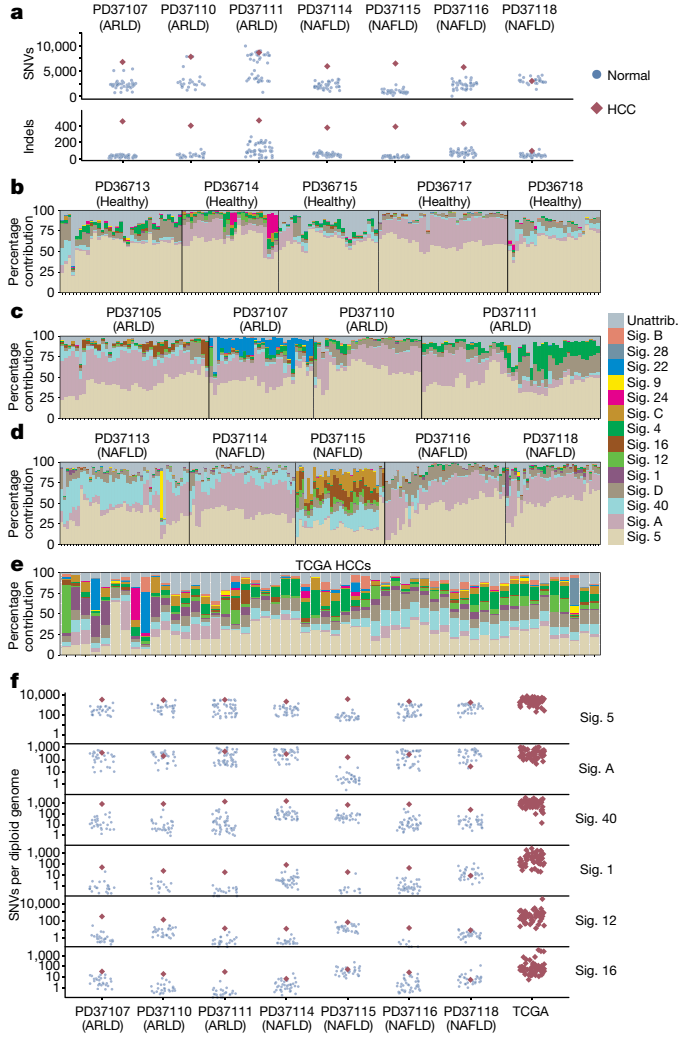



**Fig. 3 | Mutational signatures in normal liver, cirrhotic liver and HCC.** **a**, Number of somatic substitutions (SNVs; sensitivity-corrected for non-cancerous samples) and indels in each non-cancer microdissection sample (blue circles) and associated synchronous HCCs (red diamonds). **b**–**e**, Estimated proportional contributions of each mutational signature to each phylogenetically defined cluster of somatic substitutions. Data were generated using a Bayesian hierarchical Dirichlet process. Unattrib., unattributed. Stacked bar plots show proportional contributions of signatures in healthy individuals (**b**), patients with ARLD (**c**), patients with NAFLD (**d**) and 54 cases of HCC from TCGA[1] (**e**). **f**, Number of SNVs attributed to prevalent mutational signatures in each non-cancer microdissection sample (blue circles) and synchronous HCCs (red diamonds). Contributions for the TCGA samples are shown on the right. The *y* axis is on a logarithmic scale.

A major debate in the modelling of cancer development is whether cancers need higher rates of mutation to acquire sufficient driver mutations. We compared the mutational burden in cirrhotic liver to synchronous, clonally unrelated HCCs from seven patients. Synchronous HCCs carried, on average, 4,600 more mutations than matched cirrhotic liver (95% confidence interval, 3,600–5,500; $P < 10^{-18}$ using linear mixed-effect models; Fig. 3a). This indicates that rates of mutation increase during malignant transformation, either through cancer-specific mutational processes or through greater activity in cancers of ubiquitous mutational processes.

To assess which mutational processes are active in cirrhosis, we extracted mutational signatures from our 482 microdissections, as well as from the 7 synchronous HCCs and 54 HCC genomes from The Cancer Genome Atlas (TCGA)[1], using two independent algorithms (Fig. 3b–e, Extended Data Figs. 7, 8). Three major groups of mutational signatures emerged: first, those that are ubiquitous and similarly active across cirrhosis and HCC; second, those that are minor contributors in cirrhosis but universally more active in HCC; and third, those that are active in some patients but absent in others, including signatures that arise from exogenous stimuli.

In normal and cirrhotic liver, ubiquitous mutational signatures (5 and A) were prevalent across clones and, in combination, typically accounted for more than 75% of mutations. Signature 5 is widespread across cancers—including HCCs[2,4,20]—and accumulates linearly with age, suggesting that it arises from endogenous mutational processes. Signature A is the dominant cause of mutations in normal blood stem

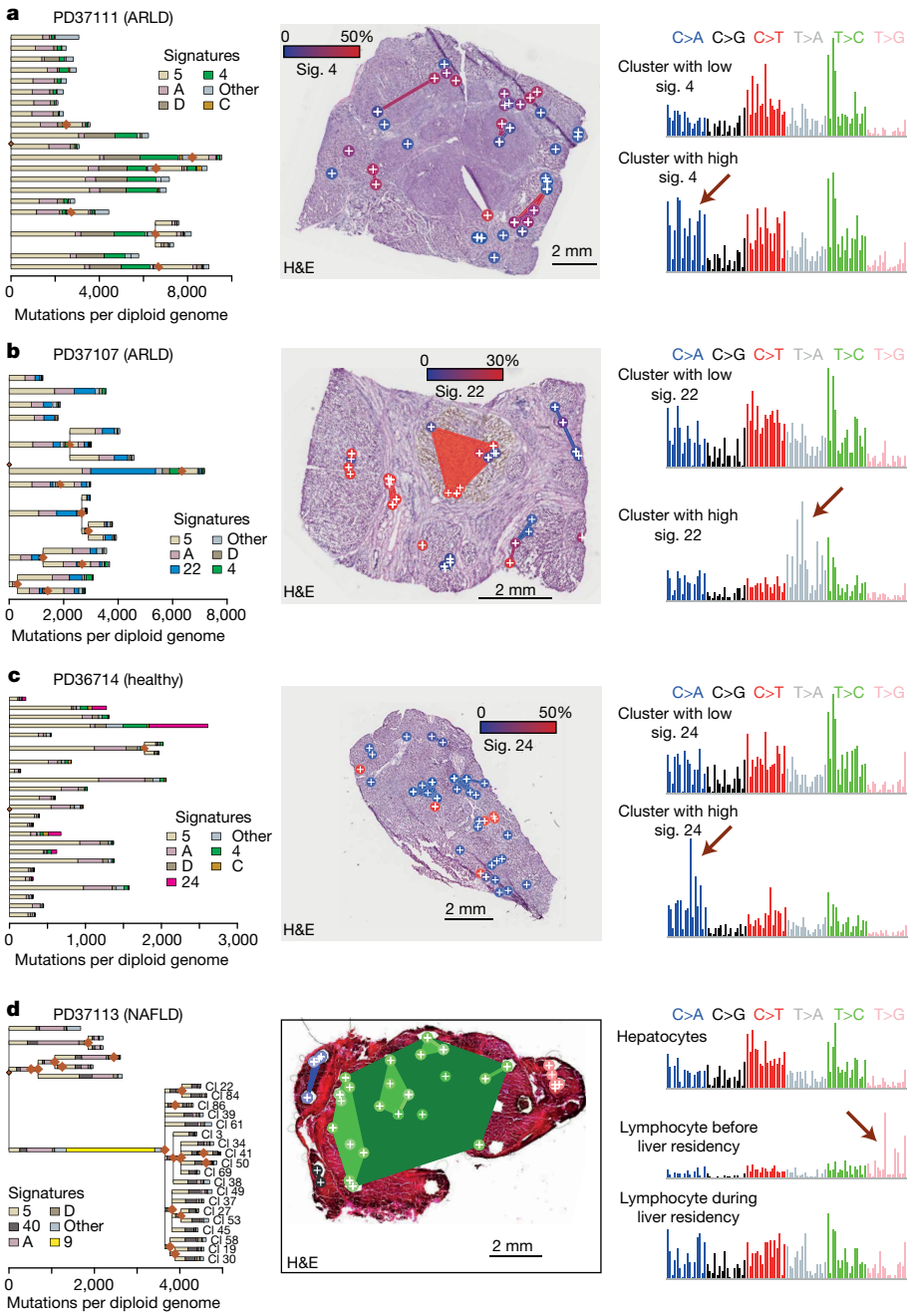



**Fig. 4 | Links between exogenous factors and mutational signatures. a**, Left, phylogenetic tree of clones in patient PD37111. Each branch is coloured by the proportion of mutations in that branch that are assigned to the different mutational signatures. Middle, overlay of the clones in **a** onto an H&E-stained liver section of patient PD37111 ($n = 39$ microdissections). Clones are coloured according to the proportion of mutations attributed to signature 4, which is linked to tobacco exposure (blue, low activity of signature 4; red, high activity of signature 4). Right, representative mutation spectra for samples with low (top) or high (bottom) burden of signature 4. The six types of substitution are labelled across the top. Within each type of substitution, the contributions from the trinucleotide context are shown as 16 bars. The 16 bars are divided into 4 sets of 4 bars, grouped by whether an A, C, G or T, respectively, is 5′ to the mutated base, and within each group of four by whether an A, C, G or T is 3′ to the mutated base. Brown arrows indicate parts of the mutation spectrum that are characteristic of the relevant mutational signatures. **b**, Overlay of mutational signatures onto phylogenetic tree of clones in patient PD37107 ($n = 41$ microdissections). The colouring of clones in the middle H&E-stained section is according to signature 22, which is linked to the carcinogen aristolochic acid. **c**, Overlay of mutational signatures onto phylogenetic tree of clones in patient PD36714 ($n = 35$ microdissections). The colouring of clones in the middle H&E-stained section is according to signature 24, which is linked to the carcinogen aflatoxin-$B_1$. **d**, Overlay of mutational signatures onto phylogenetic tree of clones in patient PD37113 ($n = 37$ microdissections). Cluster 10 has many mutations attributed to signature 9, which is linked to the process of somatic hypermutation in B lymphocytes. Cl, cluster.

cells[10,21] and leukaemias[21], which indicates that it too arises endogenously. In HCCs, although signature A accounted for a lower proportion of mutations than in normal or cirrhotic liver, the absolute numbers of mutations attributed to signature A were comparable (difference between cancer and non-cancer, 60 mutations; 95% confidence interval, −80 to 200; $P = 0.4$; Fig. 3f, Supplementary Table 10). This suggests that signature A is active in hepatocytes throughout life, but is outstripped in HCC by mutational processes that emerge during malignant transformation.

A second group of mutational signatures comprises processes that are relatively quiet in cirrhotic liver but universally more active in HCC (signatures 1, 12, 16, 40 and a new signature, D; Supplementary Table 10). One of these, signature 16, consists of T to C mutations in the ApT context and has a known transcriptional-strand bias, which includes both the preferential repair of damaged adenines on transcribed strands and increased damage on non-transcribed strands[22]. Although this signature is more active in HCCs, we do see its characteristic transcriptional-strand bias in cirrhotic liver (Extended Data Fig. 9a). Signature 1, which is caused by spontaneous deamination of methylated cytosine to thymine, is also much more active in HCC than non-malignant liver. The acceleration and universality of these signatures in HCC suggests that they reflect inbuilt DNA damage and repair processes in hepatocytes that are unmasked during malignant transformation.

The third group of mutational processes represents signatures that are seen sporadically across the cohort and that are frequently caused by exogenous factors. One, signature 4, is found in lung cancers from smokers[20], and also in HCCs, albeit with a less clear-cut relationship to tobacco[2]. Of our 14 patients, 4 had more than 10% of microdissections in which more than 5% of mutations were attributed to signature 4, demonstrating the expected transcriptional-strand bias of this signature on guanines (Extended Data Fig. 9b). Not only did signature

4 show considerable patient-to-patient heterogeneity, but there was also unexpectedly high clone-to-clone and nodule-to-nodule variability within individual livers. In one patient, for example, about half the clones we sequenced had 2,000–4,000 mutations, whereas the other half had 8,000–12,000; these differences were driven by the presence or absence of signature 4 (PD37111; Fig. 4a).

This within-patient regional variability extended to other exogenous factors. In one patient (PD37107), 20–35% of mutations were derived from signature 22 (Fig. 4b, Extended Data Fig. 9c), which is characteristic of exposure to aristolochic acid[23]. This patient grew up in Poland and spent time on holiday in Balkan states where exposure to aristolochic acid is common[24]. In a different patient (PD36714), a subset of microdissections had 10–20% of mutations that were attributable to signature 24 (Fig. 4c), which is associated with aflatoxin-$B_1$ exposure[5]. Aflatoxin-$B_1$ is produced by *Aspergillus* moulds that contaminate crops, and biomarkers of exposure to this toxin are prevalent in arable farmers[25]—the occupation of our patient. In both patients, these carcinogens showed notable variability in mutational activity over short distances, generating few mutations in some clones and hundreds to thousands in others. This regional variation in the activity of exogenous signatures is unexpected, and so far unexplained.

In one patient, we found a large clone that carried more than 2,000 mutations attributable to signature 9 (Fig. 4d)—a result of off-target somatic hypermutation in B lymphocytes[20]. A clonotypic rearrangement of *IGH* was evident, which is consistent with the notion that a single B lymphocyte subclonally diversified as it expanded in the liver (Extended Data Fig. 10). Signature 9 was only present on the ancestral trunk, whereas signatures in the subclones (acquired in the liver) were distributed in a similar manner to hepatocytes, suggesting that the hepatic microenvironment shaped the ongoing mutational processes in the lymphocytes.

In conclusion, then, non-malignant liver has considerably lower proportions of clones (less than 5%) with driver point mutations or structural variants than oesophagus or skin[11,26,27], and those present were seen in both normal and cirrhotic liver. In the cirrhotic liver, fibrosis isolated these clones, either with or without driver mutations, restricting their expansion. Moreover, driver mutations were not shared with distant synchronous HCCs, which suggests that the increased risk of cancer in chronic liver disease arises from a myriad of clones that compete independently to acquire sufficient driver mutations. Mutations in the *TERT* promoter are likely to be key events in the progression to HCC; we did not identify any *TERT* promoter mutations in cirrhotic or normal liver, but they are seen in dysplastic hepatic nodules[18,28]. The low proportion of clones with driver mutations that we observed here, and that has also been shown in exome studies performed elsewhere[29,30], means that much larger sample sizes will be needed to comprehensively map how driver mutations accumulate in the progression from normal liver through regenerative and dysplastic nodules to HCC.

These data reveal the genomic consequences of chronic liver disease—increased rates of mutation, complex structural variation (including chromothripsis and aneuploidies) and a low burden of mutations that target known HCC genes. Genomically, one middle-aged, healthy liver looks much like any other: a community of small, tightly packed clones, each comprising a few hundred cells and containing around 1,000–1,500 mutations that come from a limited palette of signatures. Unhealthy livers diverge from this norm and instead exhibit large dynasties of clones, which are sequestered by bands of fibrosis and have a repertoire of signatures that is more variable, more vigorous and more regionally variegated.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1670-9.


1. Cancer Genome Atlas Research Network. Comprehensive and integrative genomic characterization of hepatocellular carcinoma. *Cell* **169**, 1327–1341 (2017).
2. Schulze, K. et al. Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. *Nat. Genet.* **47**, 505–511 (2015).
3. Totoki, Y. et al. Trans-ancestry mutational landscape of hepatocellular carcinoma genomes. *Nat. Genet.* **46**, 1267–1273 (2014).
4. Fujimoto, A. et al. Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. *Nat. Genet.* **44**, 760–764 (2012).
5. Letouzé, E. et al. Mutational signatures reveal the dynamic interplay of risk factors and cellular processes during liver tumorigenesis. *Nat. Commun.* **8**, 1315 (2017).
6. Kan, Z. et al. Whole-genome sequencing identifies recurrent mutations in hepatocellular carcinoma. *Genome Res.* **23**, 1422–1433 (2013).
7. Guichard, C. et al. Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. *Nat. Genet.* **44**, 694–698 (2012).
8. Blokzijl, F. et al. Tissue-specific mutation accumulation in human adult stem cells during life. *Nature* **538**, 260–264 (2016).
9. Lodato, M. A. et al. Aging and neurodegeneration are associated with increased mutations in single human neurons. *Science* **359**, 555–559 (2018).
10. Lee-Six, H. et al. Population dynamics of normal human blood inferred from somatic mutations. *Nature* **561**, 473–478 (2018).
11. Martincorena, I. et al. High burden and pervasive positive selection of somatic mutations in normal human skin. *Science* **348**, 880–886 (2015).
12. Fellous, T. G. et al. Locating the stem cell niche and tracing hepatocyte lineages in human liver. *Hepatology* **49**, 1655–1663 (2009).
13. Sigal, S. H. et al. Partial hepatectomy-induced polyploidy attenuates hepatocyte replication and activates cell aging events. *Am. J. Physiol.* **276**, G1260–G1272 (1999).
14. Stephens, P. J. et al. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* **144**, 27–40 (2011).
15. Fernandez-Banet, J. et al. Decoding complex patterns of genomic rearrangement in hepatocellular carcinoma. *Genomics* **103**, 189–203 (2014).
16. Imielinski, M., Guo, G. & Meyerson, M. Insertions and deletions target lineage-defining genes in human cancers. *Cell* **168**, 460–472 (2017).
17. Martincorena, I. et al. Universal patterns of selection in cancer and somatic tissues. *Cell* **171**, 1029–1041 (2017).
18. Torrecilla, S. et al. Trunk mutational events present minimal intra- and inter-tumoral heterogeneity in hepatocellular carcinoma. *J. Hepatol.* **67**, 1222–1231 (2017).
19. Nik-Zainal, S. et al. The life history of 21 breast cancers. *Cell* **149**, 994–1007 (2012).
20. Alexandrov, L. B. et al. Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
21. Osorio, F. G. et al. Somatic mutations reveal lineage relationships and age-related mutagenesis in human hematopoiesis. *Cell Rep.* **25**, 2308–2316 (2018).
22. Haradhvala, N. J. et al. Mutational strand asymmetries in cancer genomes reveal mechanisms of DNA damage and repair. *Cell* **164**, 538–549 (2016).
23. Poon, S. L. et al. Genome-wide mutational signatures of aristolochic acid and its application as a screening tool. *Sci. Transl. Med.* **5**, 197ra101 (2013).
24. Scelo, G. et al. Variation in genomic landscape of clear cell renal cell carcinoma across Europe. *Nat. Commun.* **5**, 5135 (2014).
25. Rushing, B. R. & Selim, M. I. Aflatoxin B1: a review on metabolism, toxicity, occurrence in food, occupational exposure, and detoxification methods. *Food Chem. Toxicol.* **124**, 81–100 (2019).
26. Martincorena, I. et al. Somatic mutant clones colonize the human esophagus with age. *Science* **362**, 911–917 (2018).
27. Yokoyama, A. et al. Age-related remodelling of oesophageal epithelia by mutated cancer drivers. *Nature* **565**, 312–317 (2019).
28. Nault, J. C. et al. Telomerase reverse transcriptase promoter mutation is an early somatic genetic alteration in the transformation of premalignant nodules in hepatocellular carcinoma on cirrhosis. *Hepatology* **60**, 1983–1992 (2014).
29. Kim, S. K. et al. Comprehensive analysis of genetic aberrations linked to tumorigenesis in regenerative nodules of liver cirrhosis. *J. Gastroenterol.* **54**, 628–640 (2019).
30. Zhu, M. et al. Somatic mutations increase hepatic clonal fitness and regeneration in chronic liver disease. *Cell* **177**, 608–621 (2019).

# Methods

### Data reporting

No statistical methods were used to predetermine sample size. The experiments were not randomized and, unless otherwise stated, the investigators were not blinded to allocation during experiments and outcome assessment.

### Samples

Patients recruited at Addenbrooke's Hospital, Cambridge gave written informed consent with approval of the Local Research Ethics Committee (16/NI/0196).

Normal liver samples were obtained from patients with liver metastases from colorectal carcinoma. The liver specimens were obtained from resected liver distal to the metastases, and were confirmed to be free of tumour cells by histology. None of the patients had undergone neoadjuvant systemic therapy; one patient had undergone pre-operative portal vein embolization (PD36718) to the ipsilateral liver lobe. Liver tissue from patients with chronic liver disease was derived from explanted diseased livers at the time of transplantation. All of the patients were identified as having ARLD or NAFLD by their clinical history with the transplant hepatology and addiction psychiatry teams, as well as by explanted liver histology. None of the patients had undergone transarterial chemoembolization or other locoregional therapy on the transplant waiting list, except PD37118, who underwent a single treatment to their HCC with transarterial chemoembolization. All of the patients with chronic liver disease, except one (PD37105), demonstrated substantial pre-operative impairment of liver function as evidenced by a UK model for end-stage liver disease (UKELD) score of higher than 50.

The explant liver histology was reviewed by a specialist liver histopathologist (S.E.D.), blinded to the sequencing results. The normal liver specimens had no fibrosis and no evidence of chronic liver disease; the explanted diseased livers uniformly demonstrated cirrhosis and HCC. The background liver histology was scored according to the Kleiner system[31] on formalin-fixed paraffin-embedded (FFPE) samples away from the HCC and the fresh-frozen block used for the sequencing analysis. The Kleiner score assesses the presence of steatosis, lobular inflammation and hepatocyte ballooning to generate a cumulative NAFLD activity score (NAS). The presence or absence of cellular or nodular dysplasia was assessed globally in clinical FFPE samples (Supplementary Table 1), as well as specifically in the fresh-frozen block used for the LCM and sequencing (Supplementary Table 1). Serial H&E-stained sections from the frozen block did not demonstrate dysplasia in any of the cases (Supplementary Table 1). There was no evidence of CRC or HCC on histological review of the fresh-frozen block used for sequencing.

All tissue samples were snap-frozen in liquid nitrogen and stored at −80 °C in the Human Research Tissue Bank of the Cambridge University Hospitals NHS Foundation Trust.

### Preparation of tissue sections

Tissue biopsies were embedded in optimal cutting temperature (OCT) medium (Thermo Fisher Scientific) at −25 °C. Sections were cut at a thickness of 20 μm using a Leica cryotome and transferred onto polyethylene naphthalate (PEN) membrane slides (Thermo Fisher Scientific). For fixation, slides were treated with 70% ethanol at room temperature for 2 min. Slides were washed twice in 10% phosphate-buffered saline (PBS) at room temperature for 10 s. For staining, slides were incubated in haematoxylin for 10 s and rinsed twice in water. Slides were then incubated in eosin for 5 s and rinsed once in water. Slides were washed twice with 70% ethanol for 5 s, twice with 100% ethanol for 5 s and in xylene for 5 s. Storage was at −20 °C. Additional sections were stained for H&E, Masson's trichrome and Oil Red O by standard laboratory techniques. All slides were scanned on a Leica AT2 at ×20 magnification and a resolution of 0.5 μm per pixel.

### Laser-capture microdissection

Microdissection was performed using a laser-capture microscope (Leica Microsystems LMD 7000). For each biopsy, 48 microdissections were cut with a target size of 20,000 $\mu m^2$, which corresponds to about 400 hepatocyte cells. Images were taken before and after LCM.

### Sample lysis and DNA preparation

LCM biopsies were lysed using the Arcturus PicoPure DNA Extraction Kit (Thermo Fisher Scientific) following the manufacturer's instructions. DNA libraries for Illumina sequencing were prepared using a protocol optimized for low input amounts of DNA, as described[32].

### Whole-genome sequencing

Paired-end sequencing reads (150 bp) were generated using the Illumina X10 platform for 400 samples, resulting in a target coverage of 30×–70× per sample. To avoid the known index-hopping artefact, we chose to avoid multiplexing samples and instead sequenced one sample per flow-cell lane. To increase coverage for a subset of 96 samples, we used multiplexing and achieved 70× coverage. In addition to the LCM samples, we also sequenced a bulk sample for each biopsy and (where available) associated HCC.

The healthy liver samples came from wide resections of hepatic metastases of colorectal cancer. In each case, we sequenced the metastasis; this did not reveal any mutations that were shared between the colorectal cancer and liver, or any variants that were shared by all liver samples absent from the colorectal cancer (beyond regions of loss of heterozygosity in the cancer). Likewise, for the cirrhotic liver samples, we sequenced the matched HCC, which did not revealing any sharing of mutations. In one case, we sequenced microdissections of the fibrotic tissue, and here we also did not find mutations restricted to all liver cells.

Sequencing data were mapped to the human genome, GRCh37d5, using the BWA-MEM algorithm.

### Calling of SNVs

Substitution variants were called using the Cancer Variants through Expectation Maximization (CaVEMan) algorithm[33], using the bulk sample of the liver biopsy as the matched normal. As part of the algorithm, the variants were annotated using VAGrENT[34]. Variant calls for bulk sequencing data of the cancer samples were not further filtered. For sequencing of LCMs, post-filtering was performed in three steps.

1) Removal of duplicate counts. We noticed instances in which variant bases were counted twice, owing to the overlap of paired-end sequencing reads. We removed such double counting and re-evaluated variant calls after taking double counts into account.

2) Removal of variants that were introduced during library preparation. We noticed the presence of variants that were introduced owing to incorrect processing of cruciform DNA. Erroneous variants were often present in inverted repeats and frequently accompanied by another proximal (~1–30 bp distance). These inverted repeats can form cruciform DNA before the isolation of DNA or during library preparation. The library preparation protocol used can incorrectly process these secondary DNA structures and inadvertently introduce one or more erroneous variants. For every variant the standard deviation (s.d.) and median absolute deviation (MAD) of the variant position within the read was separately calculated for positive and negative strand reads.

In the case that the variant was supported by a low number of reads for a particular strand, the filtering was based on the statistics determined from the reads derived from the other strand. It was required that either i) 90% of supporting reads reported the variant within the first 15% of the read, as calculated from the alignment start; or ii) the MAD exceeded 0 and the s.d. exceeded 4. In the case that sufficient reads supporting the variant were available for both strands it was required for both strands separately that i) ≤90% of supporting reads reported the variant within the first 15% of the read as calculated from the alignment start; ii) the

MAD exceeded 2 and the s.d. exceeded 2; or iii) at least one strand fulfilled the criteria of a MAD greater than 1 and s.d. greater than 10.

3) Comparison with an independent panel. To remove variant calls at badly mapping sites, we compared variant calls in the sequenced samples of each donor biopsy with samples from all unrelated donors in our cohort. For each variant site we expected the reference base to be dominant and conversely expected badly mapping sites to contain frequent non-reference base counts. Thus, we counted the numbers of A, C, G, T indel calls at each variant site across all unrelated samples, resulting in a large 'pileup' table. The dominance of the reference base was evaluated at each variant site using the entropy purity metric $E$:

$$E = -\sum_i P(x_i)\ln P(x_i)$$

in which $x$ is the count of base $i \in \{A,C,G,T\}$ and the $P(x_i)$ are the fractions of base calls. Values of $E$ close to 0 indicate that almost all reads in the independent panel contain a single base. Higher values of $E$ indicate a mix of base calls at the site. To identify an optimal threshold of $E$ for the filtering of variant sites, we evaluated the entropy metric against a labelled dataset of variant calls. Specifically, during the clustering of variants using the Bayesian Dirichlet process (described below), we identified clusters that had variants with low allele frequency present in all dissections from the same donor. Manual inspection showed that such variants occurred at badly mapping sites. Thus, we labelled variant sites in those clusters as 'badly mapping' and were able to use the area under the receiver operator curve (AUC) to identify a threshold value $E_{Thr}$ of 0.16; this allowed us to separate the two labelled variant groups with an AUC of 0.99.

### Bayesian Dirichlet process for clustering VAFs across multiple samples

We extend the model previously developed for clustering VAFs of mutations called in a single sample[19] to mutation data across multiple samples from the same individual. In normal somatic cells, the vast majority of the genome retains its normal, diploid copy number, which means that we can cluster the VAFs directly (excluding mutations on the X and Y chromosomes in males). This has the considerable advantage that the Dirichlet process model we build can rely directly on conjugate prior distributions. The model includes a potential split-merge step at each cycle of the Gibbs sampler, following a previously described Metropolis–Hastings proposal for conjugate distributions[35]. The algorithm could be extended to include a correction for different copy-number states in given samples for a particular mutation through, for example, a Metropolis–Hastings update, but at considerable computational cost. The full mathematical development of the model is detailed in the Supplementary Methods.

We ran the Gibbs sampler for 15,000 iterations, dropping the first 10,000 as a burn-in. We used the ECR algorithm[36], implemented in the R package label.switching, to resolve the label-switching problem associated with mixture models. We dropped clusters that contained more than100 variant sites.

### Construction of phylogenetic trees

Phylogenetic trees were constructed manually using the pigeonhole principle, as described previously[19]. In brief, each cluster that was identified using the Bayesian Dirichlet process represented a branch of the phylogenetic tree. Nesting of trees was identified with three different levels of certainty, illustrated on a pair of branches, A and B. 1) In the case that the median VAFs of A and B exceeded 100%, the pigeonhole principle defines that A and B are nested. 2) We can assume that non-hepatocyte cells constitute a sizeable fraction of each LCM sample. Assuming a non-hepatocyte fraction of 30%, we nested branches when the VAFs of A and B exceeded 70%. This non-hepatocyte fraction was chosen as a conservative estimate of the fraction of cells intermixed in our microdissections that are not derived from the hepatocyte clone, on the basis of observed VAF peaks in our data together with single-cell RNA sequencing data from liver tissue. 3) If identical LCMs are members of both A and B, it is highly likely that A and B are nested, rather than independent branches. Thus, we also nested branches in cases in which the LCMs in one branch were a subset of the LCMs in the other (parental) branch.

For each nesting scenario, we defined the parental branch as the one with the higher median VAF in the contained LCMs. We highlighted the evidence level for nesting in each representation of phylogenetic trees, marking branches with evidence level 1 with a solid line, level 2 with a dashed line and level 3 with a dotted line.

### Analysis of driver variants

We curated a list of genes that have been found to be significantly mutated in liver cancers in a selection of published studies[1–4,6,7,37–39], as shown in Supplementary Table 5. Using the VAGrENT annotations[34], we counted any regulatory, missense, nonsense, frameshift or essential splice variant as a potential driver variant. To systematically identify genes under mutagenic selection, we used the dN/dS method[17], which screens for genes with an excess of non-synonymous mutations compared to that expected from the synonymous mutation rate.

### Sensitivity correction

We identified 138 pairs of LCMs with a midpoint-to-midpoint distance of <500 μm and at least one shared cluster according to the Bayesian Dirichlet process. These LCMs we assumed to represent the same clone, thus providing an opportunity to calculate the sensitivity of calling a variant present in one LCM in the other. If we assume the sensitivity is the same in both samples, then the maximum likelihood estimate for the sensitivity, when mutations not called in either sample are unobserved, is given by:

$$s = \frac{2n_2}{n_1 + 2n_2}$$

in which $n_2$ is the number of variants called in both LCMs in each pair and $n_1$ is the number of variants called only in one of the two LCMs. To evaluate the relationship of sensitivity with depth of coverage and VAF, we performed a logistic regression of sensitivity against these two predictors using the lm() function of the R programming language. The model fit was then used to calculate sensitivity for any LCM sample, given the coverage and VAF of the sample.

### Analysis of mutational burden

We used a linear mixed effects model to fit the number of variants per LCM sample against the disease aetiology (normal or cirrhotic) and age for each individual. We defined the ID of the individual as a random effect. The slope of the age coefficient was allowed to vary with the random effect. To facilitate the analysis, we used the lmer() function within the lme4 package of the R programming language. To determine the significance of the aetiology and age coefficients, we used an analysis of variance (ANOVA) to perform a $\chi^2$ test that compared our model with models omitting the aetiology and age coefficients, respectively.

### Targeted deep sequencing validation of mutation calls

For 96 of the microdissections sequenced by whole-genome sequencing, we performed a targeted deep sequencing validation using an Agilent RNA bait set that covered 350 recurrently mutated cancer genes. Among these genes, a total of 17 mutations were identified in the whole-genome sequencing data from the 96 samples; of these, 16 (94%) were validated, at comparable VAFs, in the targeted deep sequencing data.

### Calling of indels

Indels were called using cgpPindel[40]. Variant calls for bulk sequencing data of the cancer samples were not further filtered. To remove

artefactual calls from the LCM-derived data, we performed two post-filtering steps.

1)Assignment to SNV-based clusters. We evaluated how well the VAF distribution of each indel across the LCMs from the same donor compared with the VAF distribution of each SNV-based cluster as identified by the Bayesian Dirichlet process. Given an indel in one LCM sample, we thus counted its occurrence in all related LCMs and assigned the resulting VAF profile to the VAF profiles of the SNV clusters using a Bayes' classifier. We noticed that many indels were assigned to SNV clusters with more than 100 variants, which we had previously removed from the SNV analysis. On closer inspection we noticed that those INDELs had low VAF and occurred frequently in badly mapping regions. We thus discarded indels that were assigned to those clusters.

2) Filtering on the basis of beta-binomial overdispersion parameter. We noticed that many indels occurred with low VAF in a large number of LCMs from the same donor and were, thus, probably artefactual. To systematically identify such indels, we fitted the beta-binomial distribution to the variant counts of each indel across the LCMs from the same donor. The fitted parameter $\rho$ (the overdispersion parameter) was used to filter indel calls. A high value for parameter $\rho$ (overdispersion) occurs when some LCMs have many variant read counts and others few or none. Conversely, a low value occurs when all LCMs have a similar number of variant counts (no overdispersion). On the basis of manual inspection, we removed variant calls with $\rho < 0.02$.

### Calling of copy numbers

Copy numbers were called using the ASCAT algorithm[41], assuming an expected ploidy of 4 (to allow for physiologically polyploid hepatocytes) and 60% non-hepatocyte cell contamination for all samples. Testing of robustness around these starting points (different expected ploidy or purity values) found that the specific values used did not materially affect the output. Variant calls for bulk sequencing data of the cancer samples were not further filtered. To remove artefactual variants from the LCM-derived data, we used the SNV-based phylogenetic information. The genome was segmented into 500-bp bins and the ASCAT-based copy number of each bin was calculated. Using the binned copy-number data we calculated the median copy number in each LCM sample and ASCAT event. For each ASCAT event and LCM sample we assigned its absolute deviation from the diploid state. We compared the copy-number profile for each ASCAT event across the LCM samples with the VAF profile of each SNV cluster using cosine similarity (described below) to identify the most similar SNV cluster. Within each SNV cluster we proceeded to merge overlapping ASCAT events. Using manual inspection, we decided to keep ASCAT events if 1) they had a cosine similarity of <0.1 to an SNV cluster; and 2) their assigned SNV cluster was not removed during SNV analysis owing to having more than 100 assigned SNVs.

### Calling of structural variants

Structural variants were called using the BRASS algorithm[42] (https://github.com/cancerit/BRASS). Variant calls for bulk sequencing data of the cancer samples were not further filtered. To remove artefactual variants from the LCM-derived data, we used post-processing filters. Manual inspection of the sequencing reads identified for each structural variant showed that many reads were identical except for frameshifts at repetitive sites. We decided that such reads represented duplicates and designed a filter to systematically remove these. We removed structural variants that were supported by more than two reads after removal of duplicates. Each remaining structural variant call was manually inspected.

### Calculation of clone size

We determined the midpoint coordinates of each LCM manually from the microscopy images collected during dissection. For each LCM that belonged to a clone as determined by the Bayesian Dirichlet process, we used the chull function of the R programming language to identify the coordinates of the convex hull that included all LCMs. We identified the midpoint of each polygon as the average coordinate of all convex hull vertices. The size of the clone was then assigned to be the Euclidean distance between each convex hull vertex and the midpoint of the polygon. For clones that only consisted of a single LCM, we assigned the minimum clone size discovered across all clones.

### Extraction of mutational signatures from SNV contexts using HDP

Mutational signatures were extracted using the HDP package (https://github.com/nicolaroberts/hdp), relying on the Bayesian hierarchical Dirichlet process. The units of signature extraction were mutations assigned to individual branches of the phylogenetic tree, grouped per patient, from the LCM data. In addition, to provide a comparison against signatures extracted in HCCs, we added catalogues of somatic substitutions from 54 whole genomes sequenced by TGCA, analysed using the same core algorithms as were used for the LCM data. The tool was used without defining prior signatures. As hyperparameters we set $\alpha$ and $\beta$ to 6 for the $\alpha$ clustering parameter. Extraction was started with 40 data clusters (parameter 'initcc'). The Gibbs sampler was run with 10,000 burn-in iterations (parameter 'burnin'). With a spacing of 50 iterations (parameter 'space'), 50 iterations were collected (parameter 'n'). After each Gibbs sampling iteration, three iterations of concentration parameter sampling were performed (parameter 'cpiter'). The resulting signatures were compared to published signatures[20,43] using the cosine similarity metric described below. Extracted signatures with cosine similarity >0.9 compared to a known signature from either the COSMIC[20] or PCAWG[43] catalogue of signatures were assigned the name of the known signature with the highest similarity. Extracted signatures with cosine similarity <0.9 compared to any of the known signatures were assigned new names, which were indexed with the letters A, B and C.

### Extraction of mutational signatures from SNV contexts using SigProfiler

We used SigProfiler to extract mutational signatures, relying on the non-negative matrix factorization method[44]. In particular, we report the 'Decomposed Solution' output by the SigProfiler package.

### Cosine similarity calculation

To compare two vectors A and B, cosine similarity was calculated as follows:

$$\text{Similarity} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}}$$

### Analysis of the proportion of indels and gene expression

A list of transcribed regions was retrieved from Ensembl using the BioMaRt package[45]. We identified the subset of indel and SNV variants that overlapped with the transcribed regions. The proportion of indels in comparison to the total number of indels and SNVs per gene was calculated. Gene expression was assigned using the liver dataset from the Genotype-Tissue Expression project (GTEx)[46]. To test for the relationship between gene expression and the proportion of indels, we fitted a Poisson regression using the glm function of the R programming language. We modelled the number of indels per gene against an offset of the total number of variants per gene and the expression of the gene.

### Analysis of T>C transcriptional-strand bias at transcription start sites

We performed this analysis in a similar way to a published approach[22]. In brief, we retrieved the genomic coordinates of transcription start sites of all the highly expressed genes in the liver from GTEx[46]. We tiled the 10 kb upstream and downstream of the transcription start site into 1,000-bp bins. We overlapped all T>C (transcribed) and A>G (non-transcribed)

variant calls with the tiled regions and summed the number of variants in each tile across all included genes. We also extracted the number of T and A bases in each tile. To test whether the strand bias was significant only in transcribed regions, we fitted a Poisson regression for the number of variant calls against the following predictors: strand (transcribed, non-transcribed), distance from transcription start site (0 for upstream, 1 for downstream) and aetiology (cirrhosis, no cirrhosis), and used the number of T and A bases in each tile as the offset variable.

### Analysis of C>A and T>A transcriptional-strand bias

We used the MutationalPatterns package[47] to assign the transcription state for each C>A variant. We retrieved the genomic coordinates of all transcribed regions from Ensembl using the BioMaRt package[45] and extracted the frequencies of C and G nucleotides in these regions. To test for the significance of transcriptional-strand bias, we performed a Poisson regression for the number of C>A variants in each sample and transcription strand against factor variables for the transcription strand, the patient ID and an interaction term for the two factors. We used the C and G nucleotide frequencies as an offset variable. To test for the significance of transcriptional-strand bias for a given donor, we coded the patient ID in a binary fashion: '1' for the target donor, '0' otherwise. We proceeded to test for transcriptional-strand bias of T>A variants in a similar way, using A and T nucleotide frequencies as the offset.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

Whole-genome sequencing data across the samples reported in this study have been deposited in the European Genome-Phenome Archive (https://www.ebi.ac.uk/ega/home) in the form of BAM files, with accession number EGAD00001004578. Substitution and indel calls have been deposited on Mendeley Data with the identifier https://doi.org/10.17632/ktx7jp8sch.1 ('Somatic mutations and clonal dynamics in healthy and cirrhotic human liver').

## Code availability

Single-nucleotide substitutions were called using the CaVEMan algorithm, v.1.11.2 (https://github.com/cancerit/CaVEMan). Small insertions and deletions were called using the Pindel algorithm, v.2.2.2 (https://github.com/genome/pindel). Rearrangements were called using the BRASS (breakpoint via assembly) algorithm v.5.4.1 (https://github.com/cancerit/BRASS). Miscellaneous scripts for downstream analysis are available on Github (https://github.com/sfbrunner/liver-pub-repo). The analysis of mutational signatures was performed using the HDP hierarchical Dirichlet process package v.0.1.5, which is available on Github (https://github.com/nicolaroberts/hdp).


31. Kleiner, D. E. et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* **41**, 1313–1321 (2005).
32. Lee-Six, H. et al. The landscape of somatic mutation in normal colorectal epithelial cells. *Nature* https://doi.org/10.1038/s41586-019-1672-7 (2019).
33. Jones, D. et al. cgpCaVEManWrapper: Simple execution of CaVEMan in order to detect somatic single nucleotide variants in NGS data. *Curr. Protoc. Bioinformatics* **56**, 15.10.1–15.10.18 (2016).
34. Menzies, A. et al. VAGrENT: Variation Annotation Generator. *Curr. Protoc. Bioinformatics* **52**, 15.8.1–15.8.11 (2015).
35. Dahl, D. B. *An Improved Merge-Split Sampler for Conjugate Dirichlet Process Mixture Models. Technical Report No*. 1086 (Univ. Wisconsin–Madison, 2003).
36. Papastamoulis, P. label.switching: An R package for dealing with the label switching problem in MCMC outputs. *J. Stat. Softw*. **69**, https://doi.org/10.18637/jss.v069.c01 (2016).
37. Fujimoto, A. et al. Whole-genome mutational landscape of liver cancers displaying biliary phenotype reveals hepatitis impact and molecular diversity. *Nat. Commun*. **6**, 6120 (2015).
38. Cleary, S. P. et al. Identification of driver genes in hepatocellular carcinoma by exome sequencing. *Hepatology* **58**, 1693–1702 (2013).
39. Ahn, S.-M. et al. Genomic portrait of resectable hepatocellular carcinomas: implications of *RB1* and *FGF19* aberrations for patient stratification. *Hepatology* **60**, 1972–1982 (2014).
40. Raine, K. M. et al. cgpPindel: Identifying somatically acquired insertion and deletion events from paired end sequencing. *Curr. Protoc. Bioinformatics* **52**, 15.7.1–15.7.12 (2015).
41. Raine, K. M. et al. ascatNgs: Identifying somatically acquired copy-number alterations from whole-genome sequencing data. *Curr. Protoc. Bioinformatics* **56**, 15.9.1–15.9.17 (2016).
42. Campbell, P. J. et al. Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing. *Nat. Genet*. **40**, 722–729 (2008).
43. Alexandrov, L. et al. The repertoire of mutational signatures in human cancer. Preprint at https://www.biorxiv.org/content/10.1101/322859v2 (2019).
44. Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Campbell, P. J. & Stratton, M. R. Deciphering signatures of mutational processes operative in human cancer. *Cell Rep*. **3**, 246–259 (2013).
45. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc*. **4**, 1184–1191 (2009).
46. GTEx Consortium. Genetic effects on gene expression across human tissues. *Nature* **550**, 204–213 (2017).
47. Blokzijl, F., Janssen, R., van Boxtel, R. & Cuppen, E. MutationalPatterns: comprehensive genome-wide analysis of mutational processes. *Genome Med*. **10**, 33 (2018).



**Acknowledgements** This work was supported by a Wellcome Trust and Cancer Research UK (CRUK) Grand Challenge Award (C98/A24032). P.J.C. is a Wellcome Trust Senior Clinical Fellow (WT088340MA); S.F.B. was supported by the Swiss National Science Foundation (P2SKP3-171753 and P400PB-180790); M.A.S. is supported by a Rubicon fellowship from NWO (019.153LW.038); the Cambridge Human Research Tissue Bank is supported by the NIHR Cambridge Biomedical Research Centre; and M.H. is supported by a CRUK Clinician Scientist Fellowship (C52489/A19924).


**Author contributions** P.J.C., M.H. and S.F.B. designed the experiments; S.F.B. performed the LCM, data curation and statistical analysis, with L.A.W., M.A.S., F.A. and I.M. providing assistance and advice; M.H., S.J.A. and S.E.D. collated and analysed the clinical and histological data from the patients; N.D.R. developed the hierarchical Dirichlet process for extracting mutational signatures; L.M. and P.E. developed the LCM, DNA extraction and library production protocol used; C.A. and Y.H. assisted with sample preparation, processing and tracking; P.J.C., I.M. and M.R.S. oversaw the analysis of mutational signatures and selection analyses; P.J.C., M.H. and S.F.B. wrote the manuscript, with contributions from all authors.

**Competing interests** The authors declare no competing interests.

**Additional information**

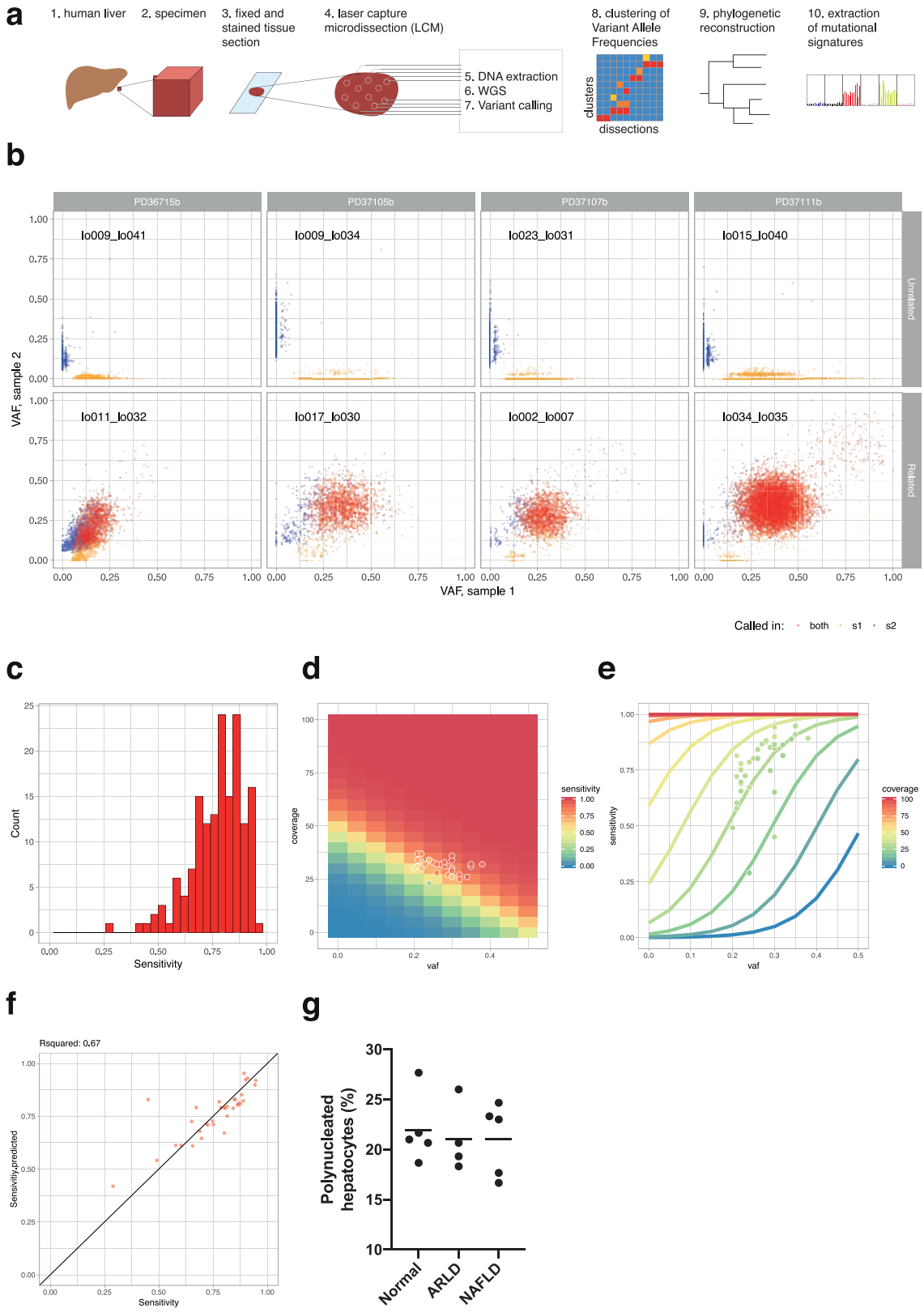

**Supplementary information** is available for this paper at https://doi.org/10.1038/s41586-019-1670-9.



**Correspondence and requests for materials** should be addressed to M.H. or P.J.C.

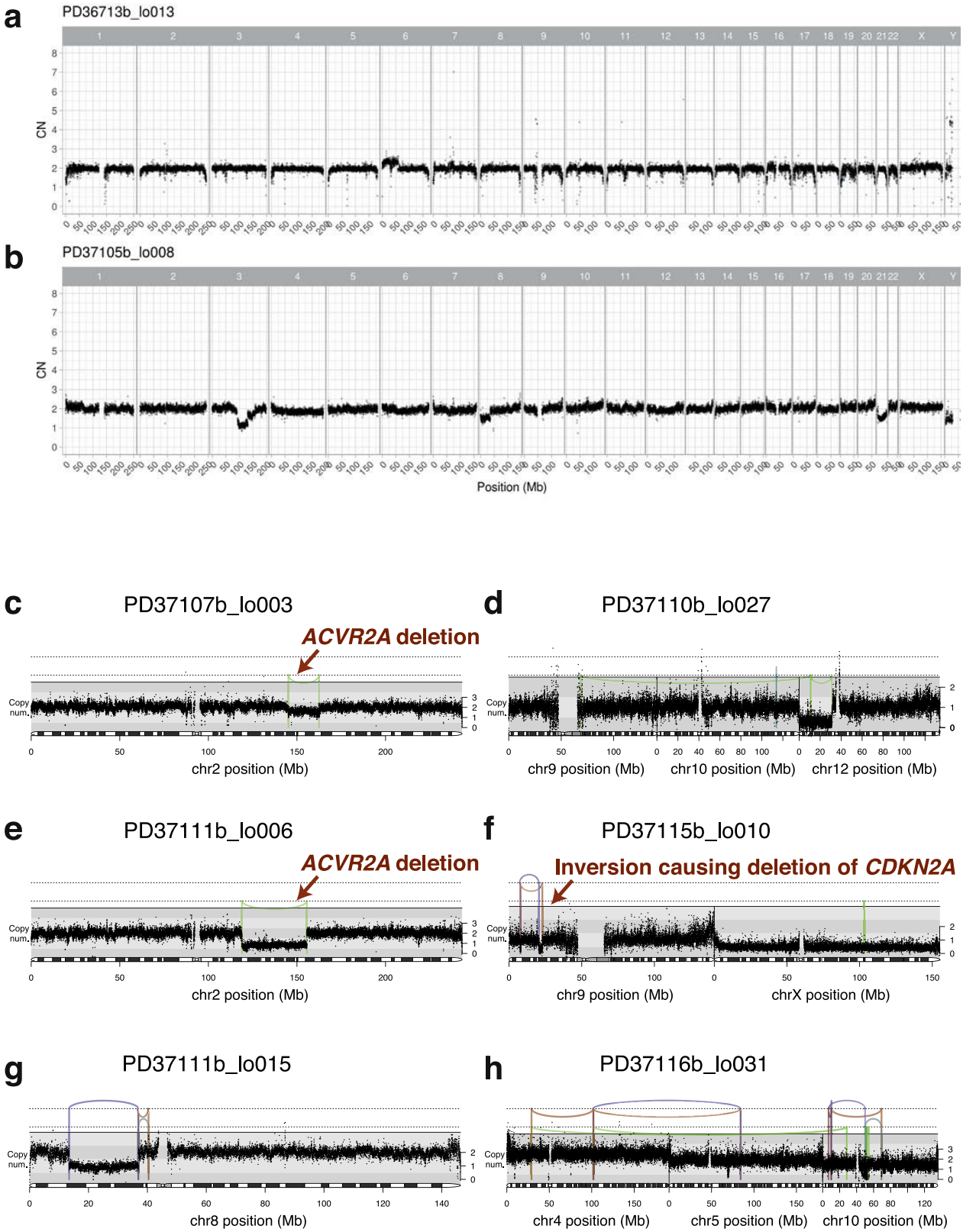


**Peer review information** *Nature* thanks Jessica Zucman-Rossi and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

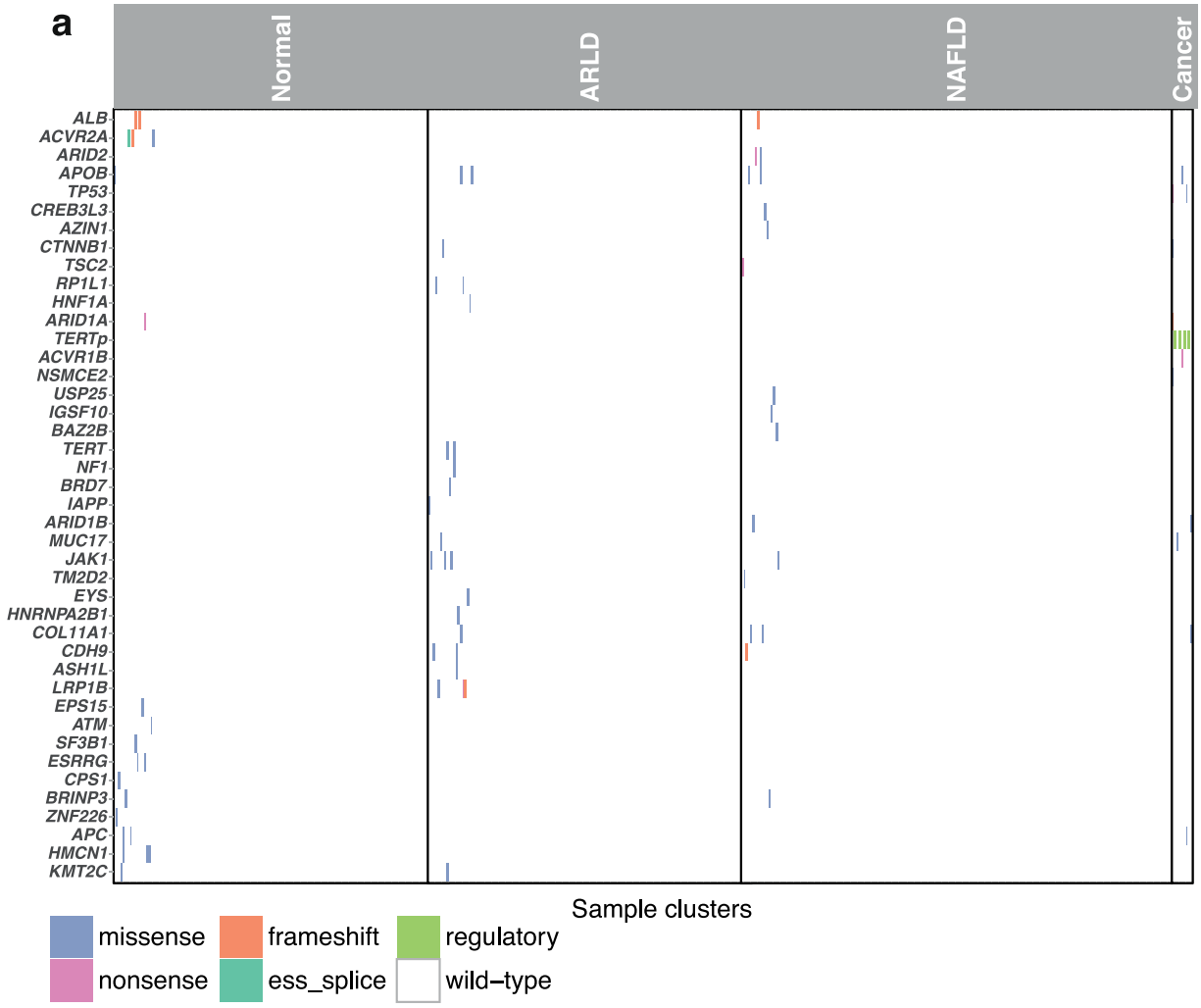



**Extended Data Fig. 1** | See next page for caption.

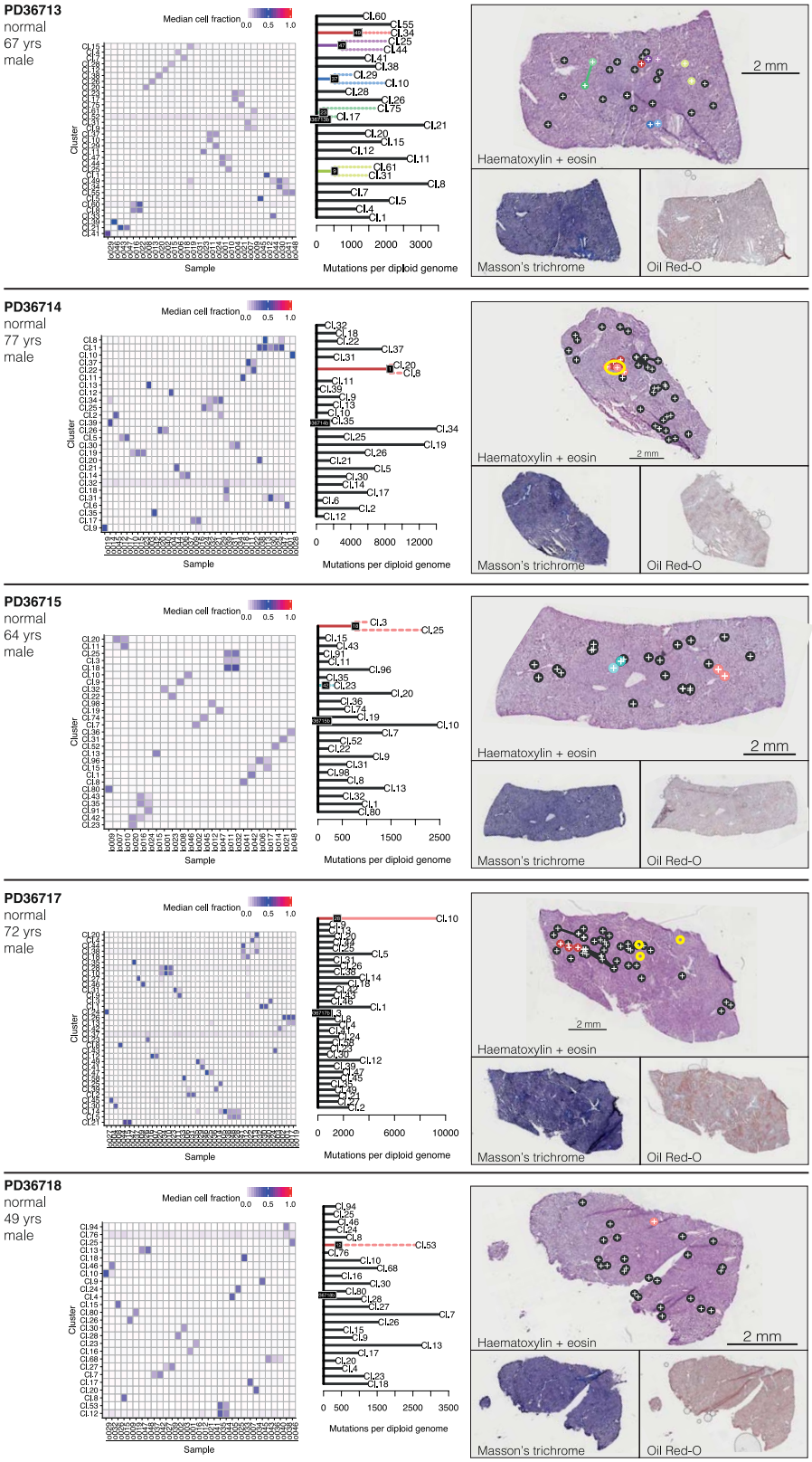
**Extended Data Fig. 1 | Sensitivity analysis of SNV calls. a**, Overview schematic of the experimental and analytical approach. **b**, Examples of the VAFs of variants from unrelated (top) and related (bottom) pairs of microdissection samples from four donors (left to right). The *x* axis represents the VAF of sample 1 from each pair and the *y* axis represents the VAF of sample 2. Each dot represents one variant. Red, variants called in both samples; yellow, variants called in sample 1; blue, variants called in sample 2. **c**, Histogram of sensitivities calculated for each sample pair. **d**, Heat map of modelled sensitivity at different values of VAF and coverage. The overlaid dots represent the sample pairs that were used to fit the model. **e**, Relationship of VAF, sensitivity and coverage according to the fitted model of sensitivity. The overlaid dots represent the sample pairs that were used to fit the model. **f**, Comparison of calculated (*x* axis) and fitted (*y* axis) sensitivity for each sample pair ($n = 34$ pairs of samples). The $R^2$ value is the Pearson's correlation coefficient. **g**, Proportion of hepatocytes that are multinucleated in the samples analysed here, estimated by counting 500 cells in each H&E-stained section ($n = 14$ patients). Each point represents the proportion for one patient in the study. The horizontal bars represent the mean for that aetiological group.




**Extended Data Fig. 2 | Copy-number and structural variants in chronic liver disease.** a, **b**, Genome-wide copy-number profiles for two samples. Black points represent the read depth of discrete windows along the chromosome, corrected to show overall copy number. Arm-level and whole-chromosome gains and losses are evident. **c**–**h**, Focal copy-number changes and structural variants. Black points represent the read depth of discrete windows along the chromosome, corrected to show overall copy number. Lines and arcs represent individual structural variants, coloured by the orientation of the joined ends (purple, tail-to-tail inverted; brown, head-to-head inverted; turquoise, tandem-duplication-type orientation; green, deletion-type orientation). Events that affect known HCC genes are marked with labelled arrows (**c**, **e**, **f**).

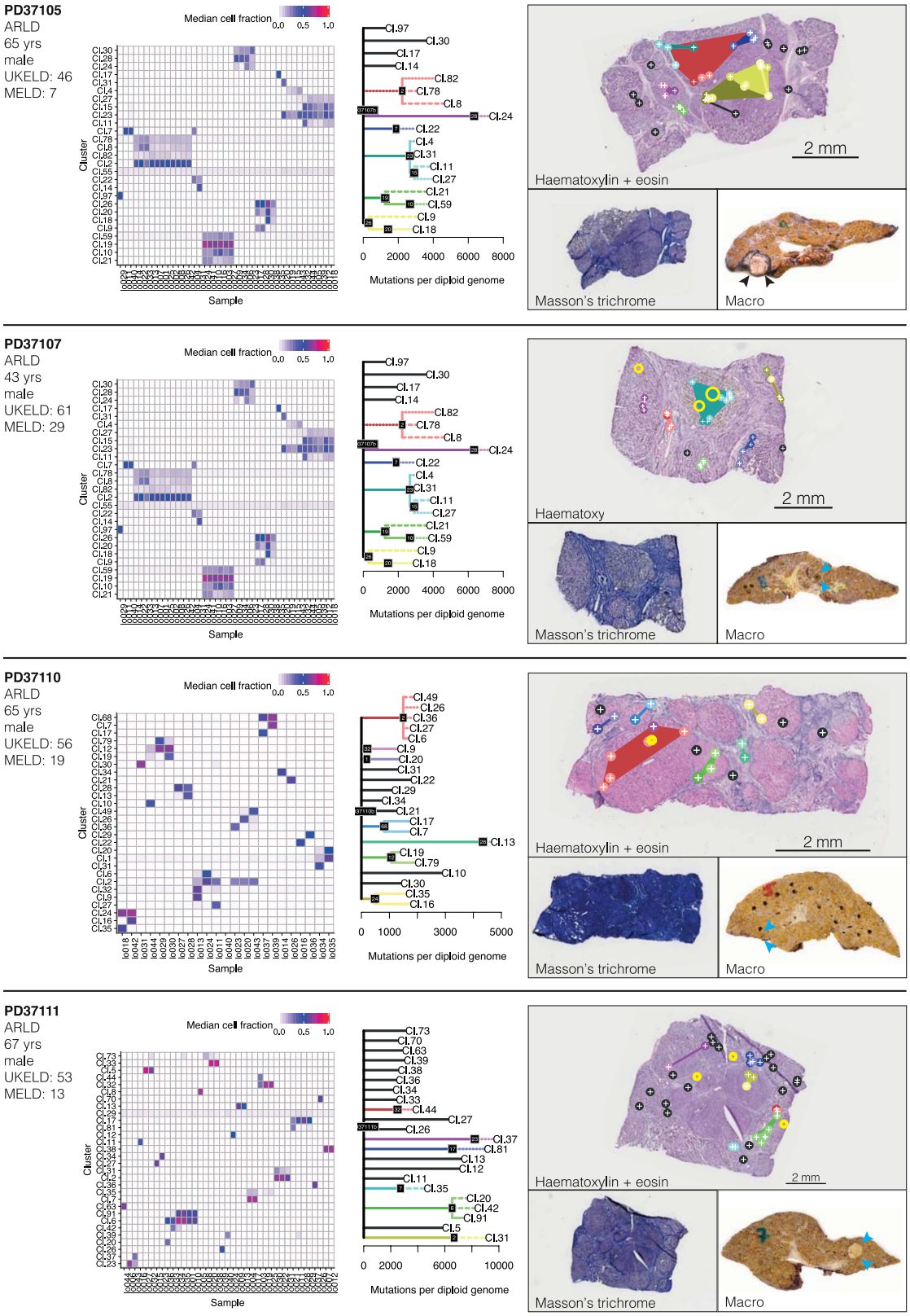



**Extended Data Fig. 3 | Events that affect known HCC genes in the cohort.** **a**, Distribution of somatic point mutations in individual microdissections (*x* axis) affecting known HCC genes (*y* axis), coloured by class of mutation according to the key underneath the panel. *TERTp*, *TERT* promoter. **b**, Genomic position of SNVs (top; light-blue strip) and indels (bottom; dark-blue strip) detected in *ALB*, the gene encoding albumin. **c**, Relationship of gene expression in liver tissue (*x* axis) and the proportion of indels as a fraction of all point mutations (*y* axis). The grey line represents a Poisson regression model with a significant (two-sided likelihood ratio test; $P < 10^{-16}$) coefficient for gene expression as a predictor for the ratio of indels ($n = 5,458$ genes included in the model). The grey ribbon represents the 99% confidence interval of the parameter estimates.

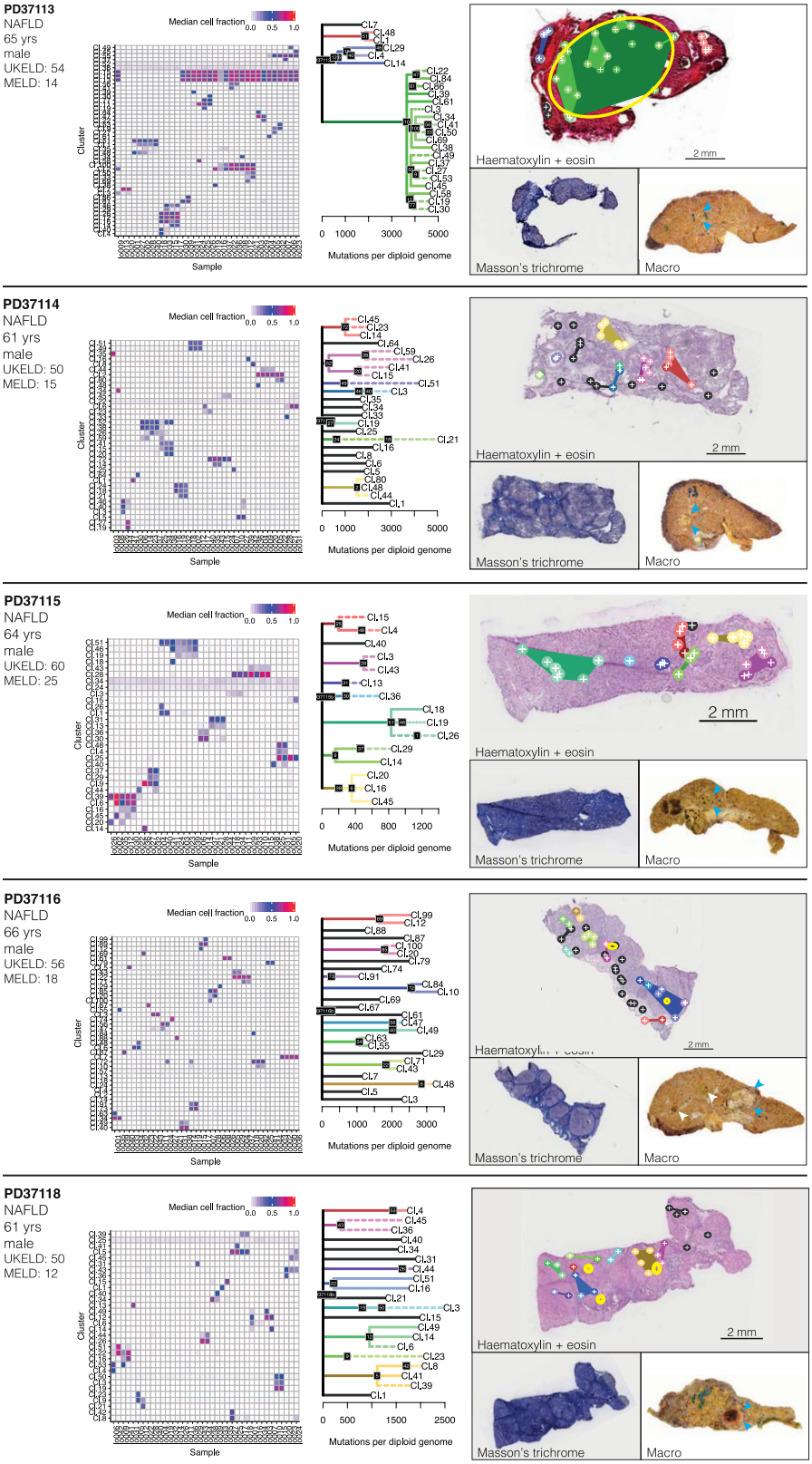



**Extended Data Fig. 4** | See next page for caption.

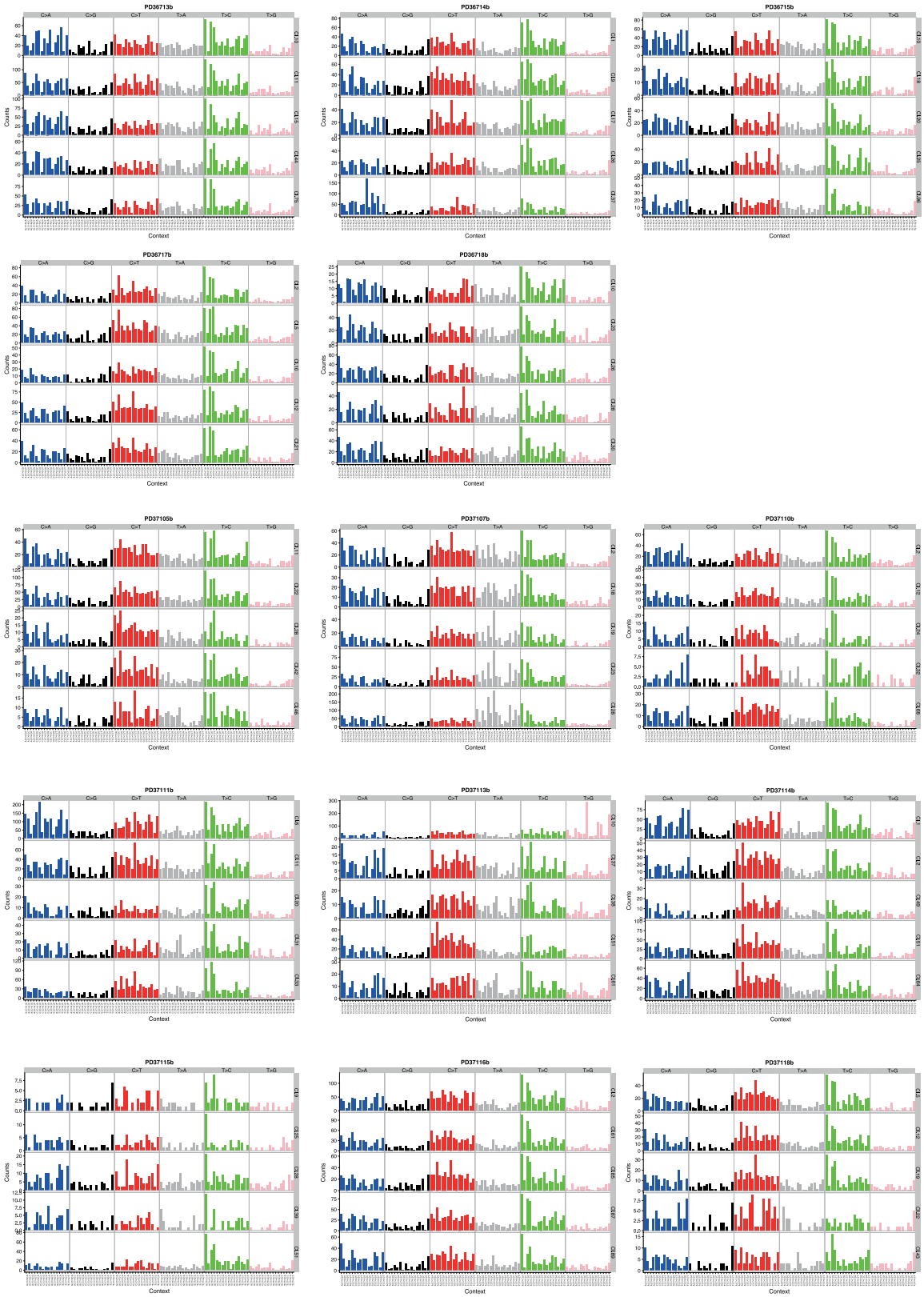
**Extended Data Fig. 4 | Phylogenetic reconstruction of hepatocyte clones in non-cirrhotic liver samples.** Left, heat maps representing the clustering of the variants observed in each microdissection sample (*x* axis) of the non-cirrhotic livers. Each cluster (*y* axis) contains mutations for which the VAFs across samples are very similar. The colour scale of the boxes represents the estimated mean VAF for that cluster in that sample. Middle, phylogenetic trees constructed from the clustering information. Solid lines indicate that nesting is in accordance with the pigeonhole principle; dashed lines indicate that nesting is in accordance with the pigeonhole principle, assuming that hepatocytes represent 70% of cells; dotted lines indicate that nesting is only based on clustering (a clone is assigned as nested if its constituent LCMs are a subset of LCMs in the parental clone). For details, see Supplementary Methods. Right, representation of clones according to the physical coordinates of the LCM samples, overlaid onto H&E-stained sections (top). Sections stained with Masson's trichrome and Oil Red O are also shown (bottom). Locations of immune or inflammatory cell infiltrates are marked with yellow rings. Sample sizes: PD36713, $n = 30$ microdissections; PD36714, $n = 35$ microdissections; PD36715, $n = 26$ microdissections; PD36717, $n = 42$ microdissections; PD36718, $n = 32$ microdissections.



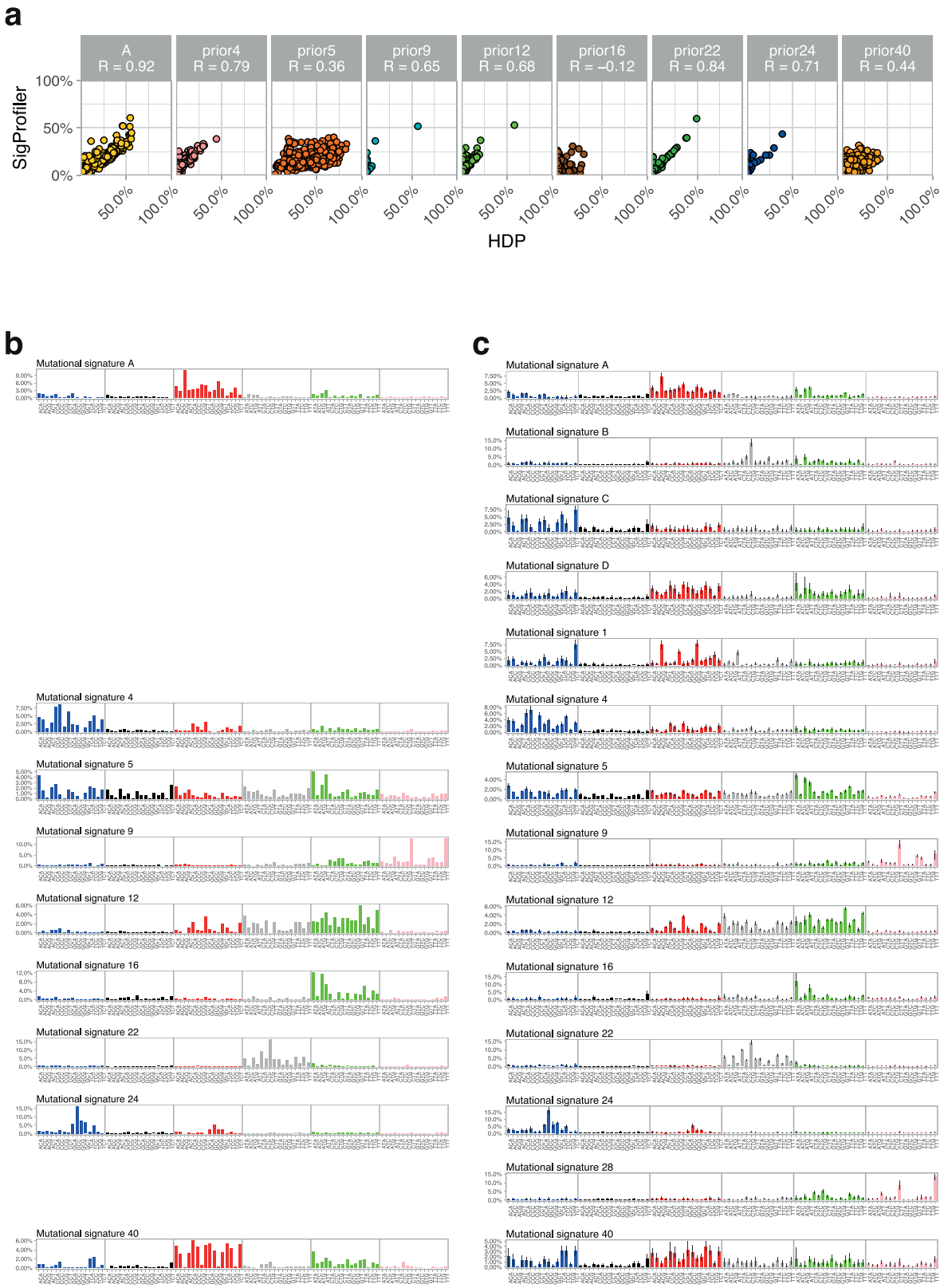

**Extended Data Fig. 5 | Phylogenetic reconstruction of hepatocyte clones in alcohol-related cirrhosis.** Analogous to Extended Data Fig. 4, but for the cirrhotic livers of donors PD37105, PD37107, PD37110 and PD37111. Right, H&E-stained sections (top); Masson's trichrome-stained sections (bottom left); and macroscopic photographs of the liver, with HCCs indicated by arrows (bottom right). Locations of immune or inflammatory cell infiltrates are marked with yellow rings. Sample sizes: PD37105, $n = 31$ microdissections; PD37107, $n = 41$ microdissections; PD37110, $n = 22$ microdissections; PD37111, $n = 39$ microdissections.



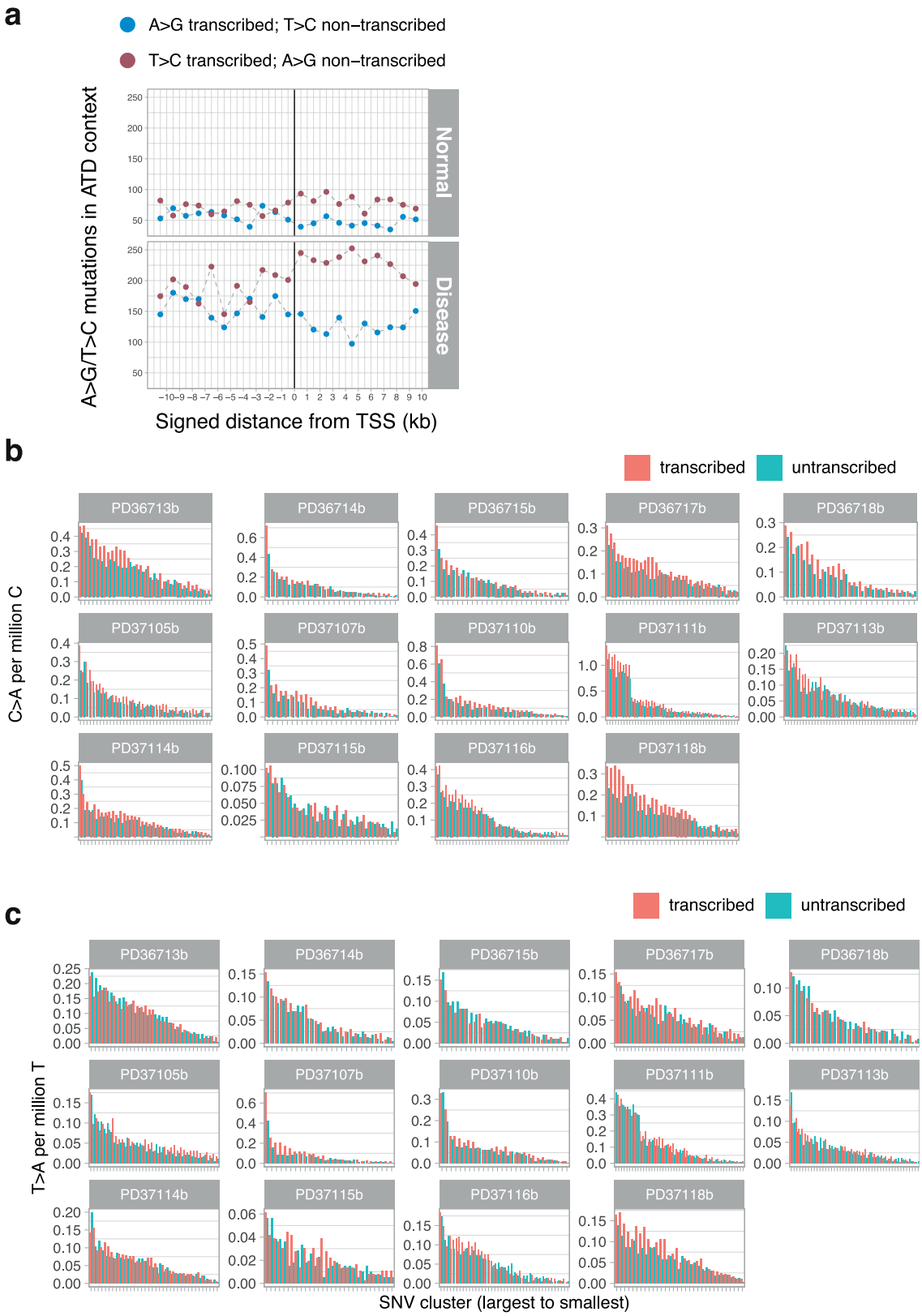

**Extended Data Fig. 6 | Phylogenetic reconstruction of hepatocyte clones in non-alcoholic fatty liver disease with cirrhosis.** Analogous to Extended Data Fig. 4, but for the cirrhotic livers of donors PD37113, PD37114, PD37115, PD37116 and PD37118. Right, H&E-stained sections (top); Masson's trichrome-stained sections (bottom left); and macroscopic photographs of the liver, with HCCs indicated by arrows (bottom right). Locations of immune or inflammatory cell infiltrates are marked with yellow rings. Sample sizes, PD37113, $n = 37$ microdissections; PD37114, $n = 41$ microdissections; PD37115, $n = 34$ microdissections; PD37116, $n = 43$ microdissections; PD37118, $n = 26$ microdissections.



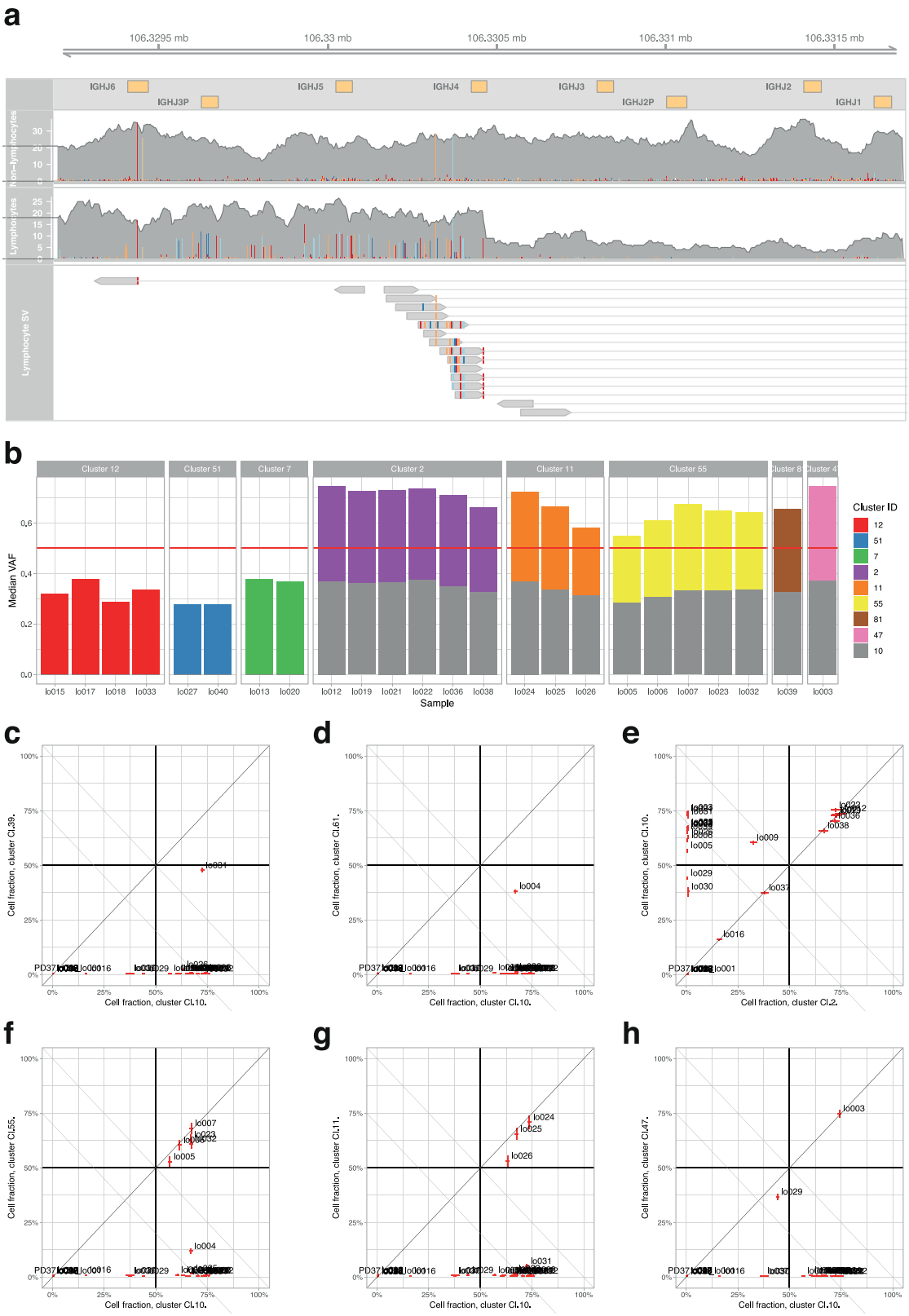

**Extended Data Fig. 7 | Mutation spectra for individual microdissections.** From each donor, we chose five clones to represent the heterogeneity in mutation spectra in the trinucleotide context. The six types of substitution are labelled across the top. Within each panel, the contributions from the trinucleotide context (bases immediately 5′ and 3′ of the mutated base) are shown.




**Extended Data Fig. 8 | Details of the extraction of mutational signatures.** **a**, Dot plots showing the concordance for signature attributions between the two signature algorithms ($n = 479$ microdissections). Mutational signatures on the *y* axis were extracted using non-negative matrix factorization and those on the *x* axis were extracted using a Bayesian hierarchical Dirichlet process. The *R* values are Pearson's correlation coefficients. **b**, Signatures extracted by non-negative matrix factorization. The six substitution classes are separated by grey vertical lines, and are presented in the following order: C>A, C>G, C>T, T>A, T>C, T>G. Within each class of mutation, the contributions from the trinucleotide context (bases immediately 5′ and 3′ of the mutated base) are shown. **c**, Signatures extracted by the Bayesian hierarchical Dirichlet process, as for **b**. Where a signature matches one from **b**, it is shown on the same row.




**Extended Data Fig. 9 | Transcriptional-strand bias in patterns of mutations.** **a**, Transcriptional-strand bias of T>C mutations at the ATD context before and after the transcription start site (TSS) of highly expressed liver genes. **b**, Bar plots representing the numbers of C>A variants on the transcribed and non-transcribed strands. Each hepatocyte clone is represented individually (*x* axis). Note the strand bias in the highly mutated clones of PD37111, in which the tobacco signature is most active; the strand bias indicates that the damaged base is the guanine, as expected for polycyclic aromatic hydrocarbons. **c**, Bar plots representing the numbers of T>A variants on the transcribed and non-transcribed strands. Each hepatocyte clone is represented individually (*x* axis). Note the strand bias in the highly mutated clones of PD37107, in which the aristolochic acid signature is most active; the strand bias indicates that the damaged base is the adenine.




**Extended Data Fig. 10** | See next page for caption.

**Extended Data Fig. 10 | Mutations in a B lymphocyte clone in a cirrhotic liver.** **a**, Illustration of a portion of the B cell receptor (*IGH*) region on chromosome 14. Shown are the coverage tracks of an LCM sample that does not belong to the lymphocyte lineage (top) and a sample that belongs to the lymphocyte lineage (middle). In the centre of the displayed region there is a drop of copy number in the lymphocyte track, which indicates a structural rearrangement. The bottom track shows the paired-end reads that contribute to a rearrangement event in the lymphocyte sample, colocalized with the drop in copy number. **b**, Application of the pigeonhole principle: if two clusters of heterozygous mutations in regions of diploid copy number are in different cells, then their median VAFs must sum to ≤0.5 (if they sum to >0.5, equivalent to a combined cellular fraction of >1, then there must be some cells that carry both sets of mutations—hence one cluster would have a subclonal relationship with the other). Cluster 10 is the cluster with the unique VDJ rearrangement of *IGH* that is shown in **a** and the large number of mutations attributed to signature 9. Clearly, samples from clusters 2, 11, 55 and so on have VAFs which, when combined with cluster 10, sum to >0.5. Therefore, they must be subclonal to cluster 10, even though they do show signature 9. **c**–**h**, Representative pairwise decision graphs for clusters of mutations. The median cellular fraction is shown for pairs of clusters across every sample from the patient. Where at least one sample falls above or to the right of the $x + y = 1$ diagonal line, those two clusters must share a nested clonal–subclonal relationship.

nature research

Corresponding author(s): Peter Campbell

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

## Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☐ | ☒ | A description of all covariates tested |
| ☐ | ☒ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ | For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☐ | ☒ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☐ | ☒ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒ | ☐ | Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |
| ☒ | ☐ | Clearly defined error bars<br>*State explicitly what error bars represent (e.g. SD, SE, CI)* |

*Our web collection on statistics for biologists may be useful.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | Image processing from sequencing data using the proprietary Illumina X10 software that is maintained, installed and distributed by Illumnina with their X10 platform. |
| Data analysis | Single-nucleotide substitutions were called using the CaVEMan (cancer variants through expectation maximization) algorithm, version 1.11.2 (https://github.com/cancerit/CaVEMan). Small insertions and deletions were called using the Pindel algorithm, version 2.2.2 (https://github.com/genome/pindel). Rearrangements were called using the BRASS (breakpoint via assembly) algorithm version 5.4.1 (https://github.com/cancerit/BRASS). Miscellaneous scripts for downstream analysis are available on Github (https://github.com/sfbrunner/liver-pub-repo). Mutational signatures analysis performed using the HDP hierarchical Dirichlet Process package version 0.1.5, available on Github (https://github.com/nicolaroberts/hdp). |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Sequence data that support the findings of this study have been deposited in the European Genome-Phenome Archive (https://www.ebi.ac.uk/ega/home) with accession number EGAD00001004578.

# Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Sample size | No formal sample size calculation was performed. Sample size was chosen to give good representation of inter-patient and intra-patient variability in mutation burden, requiring a nested design - 4-5 patients for each of 3 aetiologies (normal, alcohol-related liver disease and non-alcoholic fatty liver disease), with 30-50 microdissections per patient to enable linear mixed models to estimate both within-patient and between-patient variance. |
| Data exclusions | No data exclusions |
| Replication | Replication for sequencing and mutation calling was achieved by microdissecting the same x,y regions from adjacent z sections separated by 20 micrometres. Microdissections collected in this was were independently processed, sequenced and variant called. This approach successfully replicated the mutation calls, with quantitative results from the replication described in the manuscript and displayed in Extended Data Figure 1. |
| Randomization | Not applicable - this is a descriptive study, not an intervention study. |
| Blinding | Not applicable - all dependent variables were computationally generated (mutation counts, signatures etc) and statistical analyses were pre-specified. |

# Reporting for specific materials, systems and methods

### Materials & experimental systems

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ Unique biological materials |
| ☒ | ☐ Antibodies |
| ☒ | ☐ Eukaryotic cell lines |
| ☒ | ☐ Palaeontology |
| ☒ | ☐ Animals and other organisms |
| ☐ | ☒ Human research participants |

### Methods

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ ChIP-seq |
| ☒ | ☐ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |

## Human research participants

Policy information about studies involving human research participants

| | |
|---|---|
| Population characteristics | We analysed liver biopsies from 5 normal livers and 9 patients with cirrhosis. Clinical characteristics of the cohort are described in Supplementary Table 1. Normal livers were broadly age-matched to patients with cirrhosis, encompassing middle aged to older individuals (43-77 years). In keeping with the demographic distribution of chronic liver disease, the majority of patients were male. |

| | |
|---|---|
| Recruitment | Recruited through the Tissue Bank at Addenbrooke's Hospital, Cambridge, UK. We explicitly studied samples from patients with chronic liver disease who had a synchronous hepatocellular carcinoma, meaning that patients were at advanced stages of disease. Otherwise, we anticipate no recruitment biases affecting the sample mix. |

Article

# The microbiota regulate neuronal function and fear extinction learning

https://doi.org/10.1038/s41586-019-1644-y

Received: 23 August 2018

Accepted: 5 September 2019

Published online: 23 October 2019

Coco Chu[1], Mitchell H. Murdock[2,3,4], Deqiang Jing[3,4,5], Tae Hyung Won[6], Hattie Chung[7], Adam M. Kressel[8,9,10], Tea Tsaava[8], Meghan E. Addorisio[8], Gregory G. Putzel[1], Lei Zhou[1], Nicholas J. Bessman[1], Ruirong Yang[3,4,5], Saya Moriyama[1], Christopher N. Parkhurst[1], Anfei Li[3,4], Heidi C. Meyer[3], Fei Teng[1], Sangeeta S. Chavan[8,9,11], Kevin J. Tracey[8,9,11], Aviv Regev[7,12], Frank C. Schroeder[6], Francis S. Lee[3,4,5], Conor Liston[2,3,4]* & David Artis[1,13]*

Multicellular organisms have co-evolved with complex consortia of viruses, bacteria, fungi and parasites, collectively referred to as the microbiota[1]. In mammals, changes in the composition of the microbiota can influence many physiologic processes (including development, metabolism and immune cell function) and are associated with susceptibility to multiple diseases[2]. Alterations in the microbiota can also modulate host behaviours—such as social activity, stress, and anxiety-related responses—that are linked to diverse neuropsychiatric disorders[3]. However, the mechanisms by which the microbiota influence neuronal activity and host behaviour remain poorly defined. Here we show that manipulation of the microbiota in antibiotic-treated or germ-free adult mice results in significant deficits in fear extinction learning. Single-nucleus RNA sequencing of the medial prefrontal cortex of the brain revealed significant alterations in gene expression in excitatory neurons, glia and other cell types. Transcranial two-photon imaging showed that deficits in extinction learning after manipulation of the microbiota in adult mice were associated with defective learning-related remodelling of postsynaptic dendritic spines and reduced activity in cue-encoding neurons in the medial prefrontal cortex. In addition, selective re-establishment of the microbiota revealed a limited neonatal developmental window in which microbiota-derived signals can restore normal extinction learning in adulthood. Finally, unbiased metabolomic analysis identified four metabolites that were significantly downregulated in germ-free mice and have been reported to be related to neuropsychiatric disorders in humans and mouse models, suggesting that microbiota-derived compounds may directly affect brain function and behaviour. Together, these data indicate that fear extinction learning requires microbiota-derived signals both during early postnatal neurodevelopment and in adult mice, with implications for our understanding of how diet, infection, and lifestyle influence brain health and subsequent susceptibility to neuropsychiatric disorders.

Pavlovian fear conditioning is an evolutionarily conserved associative learning process that is critical for the survival of an organism and for the ability to respond appropriately to neutral stimuli that reliably predict dangerous or aversive outcomes[4]. In the classical fear conditioning paradigm, extinction learning occurs when repeated cue presentations are no longer paired with an unconditioned stimulus (such as a foot shock) and the organism learns to modify its behaviour accordingly. Deficits in extinction learning after an environmental threat has passed have been implicated in multiple neuropsychiatric disorders, including post-traumatic stress disorder and other anxiety disorders[5]. Clinical and epidemiological studies have reported correlations between changes in the microbiota and other neuropsychiatric disorders[6–8]. Animal studies indicate that the absence or modification of the intestinal microbiota affects neurogenesis[9], cortical myelination[10], the function of the blood–brain barrier[11] and maturation of microglia[12], as well as social behaviour, stress-related responses and fear learning[3,13,14]. However, there are conflicting reports on how the microbiota influence behaviour[13–16], and the mechanisms through which the microbiota regulate associative learning and its neurobiological substrates remain unclear.

## Lack of microbiota impairs extinction learning

To test whether the microbiota influence fear conditioning and extinction, we first treated adult mice with antibiotics (termed ABX mice)[17] and used a classical cued fear conditioning and extinction learning paradigm[18]. ABX mice and control mice showed comparable food and

A list of affiliations appears at the end of the paper.

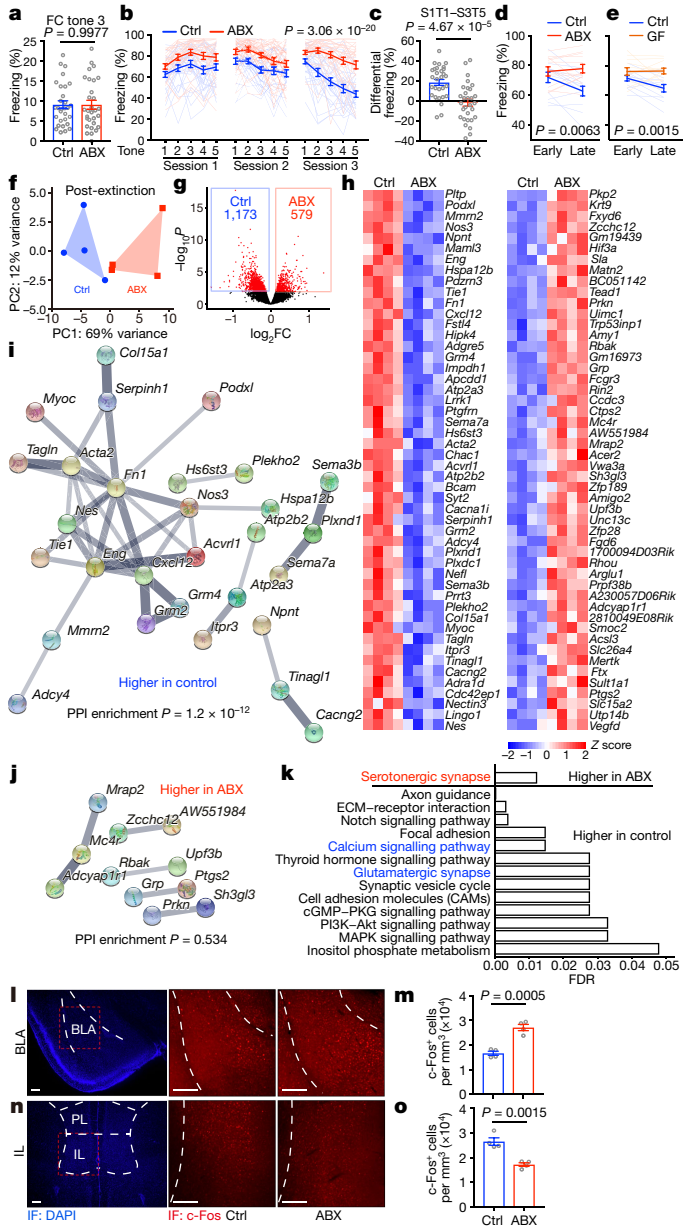



**Fig. 1 | ABX and GF mice are less prone to fear extinction. a–c**, Acquisition of fear conditioning (FC) (**a**), fear extinction over the course of three days or sessions (**b**) and fear conditioning after three days (**c**) in control (Ctrl) and ABX mice. S, session; T, tone. Data pooled from two independent experiments. $n = 30$ mice per group. Mean ± s.e.m.; unpaired two-sided $t$-test (**a**, **c**); for **b**, area under the curve (AUC) was calculated for each mouse within each group followed by unpaired two-sided $t$-test between groups. **d**, **e**, Fear extinction in control versus ABX (**d**) or control versus GF (**e**) mice in the single-session 30-tone fear extinction assay. Data pooled from two independent experiments, $n = 12$ mice per group. Mean ± s.e.m.; the AUC was calculated for each mouse within each group followed by unpaired two-sided $t$-test between groups. **f**, Principle component analysis (PCA) of mouse mPFC transcriptome after fear extinction. $n = 4$ mice per group. Permutational multivariate analysis of variance (PERMANOVA): $F = 5.00$, Df = 1, $P = 0.027$. **g**, Volcano plot of differential expression of genes in ABX versus control groups in **f**. Red dots, differentially expressed genes (DESeq2 Wald test, false discovery rate (FDR) < 0.1). FC, fold change. **h**, Heat maps showing the top 50 most significantly downregulated or upregulated genes in **g**. Low-expression genes with mean normalized counts in the bottom 20th percentile were excluded. **i**, **j**, STRING network visualization of the genes in **h**. Edges represent protein–protein associations. Disconnected nodes were excluded. **k**, Significantly enriched KEGG pathways based on all differentially expressed genes in **g**. **l**–**o**, Immunofluorescence (IF) staining (**l**, **n**) and the density of c-FOS$^+$ neurons (**m**, **o**) in the BLA (**l**, **m**) or IL (**n**, **o**) of control and ABX mice after fear extinction session 3. Data are representative of two independent experiments. $n = 4$ mice per group. Mean ± s.e.m.; unpaired two-sided $t$-test. PL, prelimbic. Scale bar, 200 µm.

water intake and weight gain (Extended Data Fig. 1a–c). The bacterial burden was 600-fold lower in ABX mice than in control mice (Extended Data Fig. 1d), and 16S ribosomal DNA (rDNA) sequencing revealed an antibiotic-induced shift in bacterial community structure (Extended Data Fig. 1e–g). Following fear conditioning, ABX mice displayed equivalent freezing behaviour to control mice, indicating that their acquisition of fear conditioning was normal (Fig. 1a). Extinction learning reduced conditioned freezing in control mice[18]. By contrast, extinction learning was significantly impaired in ABX mice (Fig. 1b, c). To further examine the influence of the microbiota on extinction learning, we performed a similar cued fear conditioning and extinction learning assay in adult germ-free (GF) mice. To maintain the microorganism-free status of the GF mice, we used a modified single-session fear extinction protocol[19]. Again, both ABX and GF mice exhibited impaired extinction learning (Fig. 1d, e). These data show that signals derived from the microbiota are necessary for optimal extinction of conditioned fear responses.

The vagus nerve is one route of neuronal communication between the intestine and the brain[20,21]. We investigated whether the vagus nerve is involved in extinction learning deficits following manipulation of the microbiota by carrying out surgical vagotomy on adult mice. Vagotomized ABX mice exhibited similar deficits in extinction learning to sham-operated ABX mice, suggesting that the extinction learning deficits in ABX mice are independent of the vagus nerve (Extended Data Fig. 2).

Given that microbiota-derived signals can regulate the immune system and that immune cells can influence brain function and behaviour[22–24], we tested whether the extinction learning deficits were associated with alterations in immune responses in the brain. Compared to control mice, ABX and GF mice showed no differences in the percentages and numbers of CD45$^{high}$ leukocytes in the brain (Extended Data Fig. 3a, b, e), and no differences in the percentages of CD4$^+$ T cells, CD8$^+$ T cells, CD19$^+$ B cells, CD11c$^+$ dendritic cells, F4/80$^+$ macrophages or Ly6C$^{high}$ monocytes (Extended Data Fig. 3c, d, f–j). Moreover, *Rag1*$^{-/-}$ mice, which lack adaptive immune cells, exhibited normal extinction learning (Extended Data Fig. 3k), whereas GF *Rag1*$^{-/-}$ mice showed deficits in extinction learning (Extended Data Fig. 3l), indicating that the adaptive immune system is not required for extinction learning deficits in ABX and GF mice.

Given that deficits in extinction learning appear to occur independently of changes in the immune system, we investigated their neuroanatomical basis. We performed genome-wide transcriptional profiling of the medial prefrontal cortex (mPFC), an area of the brain known to be crucial for extinction learning[25], from adult ABX and control mice. mPFC tissue dissected from adult ABX and control mice that had not undergone fear conditioning and extinction exhibited comparable transcriptomes (Extended Data Fig. 4a, b). However, extinction learning led to significant differences in the transcriptomes of mPFC tissue from ABX and control mice (Fig. 1f, g). Search tool for recurring instances of neighbouring genes (STRING) analysis depicted networks of interactions of the differentially expressed genes (DEGs) between ABX and control samples (Fig. 1h–j); Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) enrichment analyses identified pathways that are associated with neuronal activity, synapse function, CNS maturation and the regulation of synaptic plasticity and the development of postsynaptic dendritic spines (Fig. 1k, Supplementary Table 1).

To test whether alterations in gene expression associated with these neuronal processes were associated with changes in neuronal activity, we examined neuronal activity in fear learning circuits by analysing c-FOS expression[26] in the basolateral amygdala (BLA), which is critical for encoding and storing conditioned fear memory[27], and in the infralimbic region (IL) of the mPFC, which facilitates extinction learning[25].

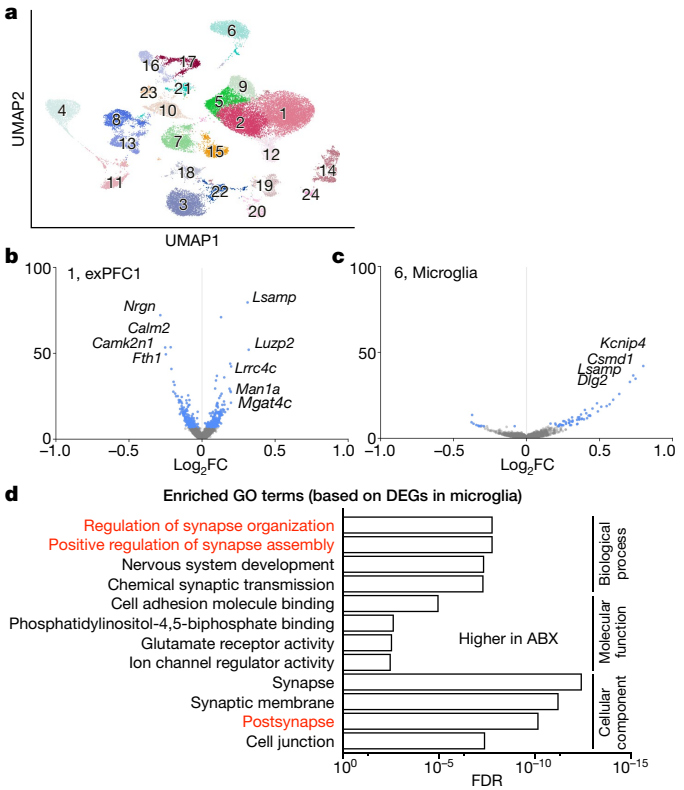



**Fig. 2 | Excitatory neurons and microglia are affected in ABX mice. a**, Single nucleus profiles coloured by cell type (see Extended Data Fig. 5 for annotations). $n = 2$ mice per group. **b**, **c**, Differentially expressed genes (blue dots: *z*-test calculated on coefficients of mixed linear model, Bonferroni-corrected $P < 10^{-7}$) of ABX versus control in excitatory neuron subset 1 (exPFC1, **b**) or 6 (microglia, **c**). **d**, GO terms that were significantly enriched among the differentially expressed genes in the microglial cluster in **c**.

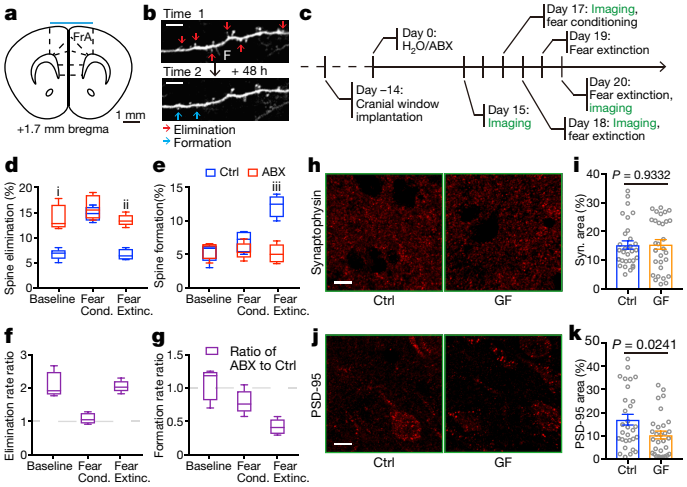



**Fig. 3 | Defective extinction-learning-related dendritic spine formation in ABX mice. a**, Diagrammatic representation of a coronal section of the mPFC showing the imaging site (cyan bar). FrA, frontal association cortex. **b**, Example images of neuronal dendritic branch segments at two consecutive imaging time points. Red arrows, spine elimination. Blue arrows, spine formation. Scale bar, 5 µm. F, filopodia. **c**, Timeline of transcranial two-photon spine imaging. **d**, **e**, Percentage of spine elimination (**d**) and formation (**e**) at baseline, during fear conditioning and during fear extinction. **f**, **g**, Ratio of ABX to control for spine elimination rate (**f**) and spine formation rate (**g**). Data in **d**–**g** are pooled from three independent experiments. $n = 5$ mice per group. Centre line, median; box, 25th and 75th percentiles; whiskers, minimum to maximum. Unpaired two-sided *t*-tests were used in **d**, **e**. i, $P = 0.0010$; ii, $P = 4.38 \times 10^{-5}$; iii, $P = 0.0002$. **h**–**k**, Immunofluorescence staining and the area of synaptophysin (**h**, **i**) or PSD-95 (**i**, **k**) in the mPFC of control and GF mice. Data are pooled from two independent experiments. $n = 6$ mice per group. Each symbol represents one region of interest (ROI); five ROIs per mouse. Mean ± s.e.m.; unpaired two-sided *t*-test. Scale bars, 10 µm.

Compared to control mice, ABX and GF mice exhibited a higher density of c-FOS$^+$ neurons in the BLA (Fig. 1l, m, Extended Data Fig. 4c, d) and lower density of c-FOS$^+$ neurons in the IL (Fig. 1n, o, Extended Data Fig. 4e, f), which is consistent with their deficits in extinction learning.

## Neuronal and glial changes in ABX mice

To define the cell subsets in the mPFC that contribute to the effect of the microbiota on extinction learning, we performed single-nucleus RNA sequencing (snRNA-seq) of mPFC samples dissected from ABX and control mice after extinction learning, and identified 24 cell clusters (Fig. 2a, Extended Data Fig. 5a). Changes in the microbiota were associated with significant changes in gene expression in multiple clusters (Extended Data Fig. 6). Among the neuronal clusters, excitatory neurons (Fig. 2b, Extended Data Fig. 6) were more affected than inhibitory neurons (including PVALB$^+$TAC1$^+$, SST$^+$, VIP$^+$ and NPY$^+$ subsets) (Extended Data Figs. 5b, 6). Astrocytes, myelinating oligodendrocytes and microglia also showed changes in gene expression (Extended Data Fig. 6). DEGs that were shared across subsets of excitatory neurons (Extended Data Fig. 7) and across multiple cell types (Extended Data Fig. 8) were linked to synapse-related pathways and calcium signalling pathways (Supplementary Tables 2, 3), which is consistent with our bulk RNA-seq data (Fig. 1k) and supports a model in which gene expression is altered in brain-resident cells—including specific cell populations such as excitatory neurons and microglia—following manipulation of the microbiota.

Given that microglia are important for maintaining neuronal function and brain health by dynamically regulating synaptic pruning and surveying their local microenvironments, and have been reported to be affected by the microbiota[12,28], we further investigated the DEGs in microglia (Fig. 2c). The microglial DEGs were enriched in pathways related to synapse organization and synapse assembly (Fig. 2d, Supplementary Table 4), suggesting that deliberate manipulation of the microbiota may alter microglia-mediated synaptic pruning. In addition, consistent with the literature[12], we found elevated percentages and numbers of microglia in GF mice, with elevated expression of CSF1R and F4/80 (Extended Data Fig. 9a–d). The percentages and numbers of microglia in ABX mice were not changed, with no changes in CSF1R expression, but F4/80 expression was elevated (Extended Data Fig. 9e–h). CSF1R and F4/80 are strongly developmentally regulated, and their expression decreases during maturation[29]. Together, these data suggest that microglia in GF and ABX mice exhibit an immature state reminiscent of developing juvenile microglia, which may in turn contribute to deficits in extinction learning by disrupting dendritic spine remodelling.

## Defective extinction-related spine remodelling

Next, we used two-photon laser scanning microscopy to directly quantify the remodelling of postsynaptic dendritic spines in the mPFC (Fig. 3a) during cued fear conditioning and extinction learning in transgenic THY1–YFP-H reporter mice, which express yellow fluorescent protein in neurons, following manipulation of the microbiota in adulthood. Postsynaptic dendritic spines are membranous protrusions on neuronal dendrites that form primarily excitatory synapses with presynaptic axonal inputs and are dynamically remodelled during learning and development[30–33]. Fear conditioning and extinction learning are correlated with opposing effects on the formation and elimination of dendritic

spines in the mPFC[34]. We acquired images of the same dendritic spines during a baseline period, and before and after fear conditioning and extinction learning (Fig. 3b, c). Compared to control mice, baseline spine elimination rates were significantly elevated in ABX mice (Fig. 3d, f), whereas baseline spine formation rates were unaffected (Fig. 3e, g). Consistent with previous findings[34], cued fear conditioning and extinction learning had opposing effects on spine remodelling in control mice. Fear conditioning increased spine elimination rates in control mice (Fig. 3d), such that there was no significant difference in spine elimination or formation rates in the 24 h after conditioning between ABX mice and control mice (Fig. 3d–g). By contrast, extinction-learning-related spine remodelling was significantly altered in ABX mice. Extinction learning increased spine formation rates in control mice but not in ABX mice (Fig. 3e, g), and spine elimination rates remained persistently elevated in ABX mice relative to control mice (Fig. 3d, f).

Consistent with elevated spine elimination in ABX mice, we observed comparable expression of the presynaptic marker synaptophysin but lower expression of the postsynaptic marker PSD-95 in the mPFC of GF mice compared to control mice (Fig. 3h–k). In addition, the expression of *Dlg4* in excitatory neuron subset 1 (exPFC1) was downregulated in samples from ABX mice compared to those from control mice. Together, these data indicate that alterations in the composition of the microbiota are associated with deficits in learning-induced spine plasticity. Notably, plasma corticosterone levels were comparable in control, ABX and GF mice[35] (Extended Data Fig. 10a, b), indicating that the function of the hypothalamic–pituitary–adrenal axis may not be altered and is not a major driver of microbiota-mediated changes in spine remodelling and fear extinction learning.

## Defective tone-encoding ensembles in ABX mice

To investigate whether signals derived from the microbiota regulate learning-related neuronal activity in the mPFC, we used two-photon imaging and a genetically encoded calcium sensor (AAV5–hSyn–GCaMP6s) to quantify neuronal activity during extinction learning in control and ABX mice (Fig. 4a–c). We identified mPFC neurons that encoded the conditioned stimulus, and identified two differentially responding functional cell types. One neuronal population (representing 13.5% of all cells) exhibited equivalently reduced activity during tone presentations in both control and ABX mice (Fig. 4d, e). By contrast, a second population (representing 14.9% of all cells) displayed increased activity during tone presentations (Fig. 4f). In this latter population, neuronal activity during tone presentations was modestly but significantly lower in ABX mice than in control mice (Fig. 4g), consistent with the deficits in spine formation and behaviour in ABX mice. Notably, 26.8% of these neurons also encoded the precise timing of the tones, exhibiting tone-locked activity that increased and decreased in response to the onset and offset of each tone, respectively (Fig. 4h). Again, tone-locked activity in these multicellular tone-sensitive ensembles was significantly reduced in ABX mice compared to control mice (Fig. 4i). In conjunction with the RNA-seq and dendritic spine remodelling analyses, these data indicate that dysbiosis of the gut microbiota disrupts learning-related spine formation and interferes with the emergence of multicellular tone-encoding ensembles.

## Diverse microbiota restores extinction learning

To test whether the extinction learning deficits caused by altered microbiota can be rescued by colonization with defined individual microorganisms or consortia of microorganisms, we performed fear conditioning and extinction learning in gnotobiotic mice colonized with bacteria that are known to influence other physiologic processes. Notably, following colonization with segmented filamentous bacterium (SFB)[36], *Clostridium* species[37], *Enterobacter* species[38] or altered Schaedler flora (ASF)[39], these gnotobiotic mice still exhibited impaired extinction learning compared to control mice (Fig. 5a), suggesting that a more diverse microbiota is required for normal extinction learning and fear extinction behaviour.

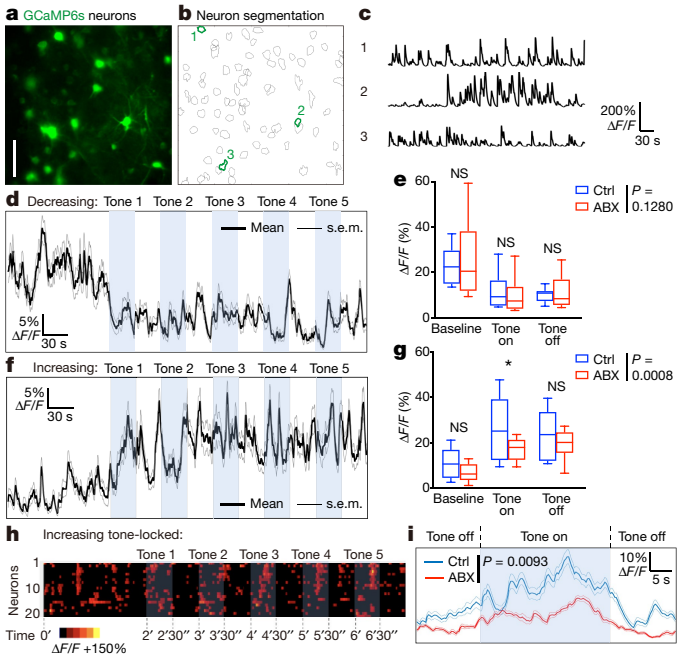



**Fig. 4 | Defective ensemble calcium dynamics in the mPFC of ABX mice.** **a**, Example false-colour image (mean projection over time) of GCaMP6s-expressing neurons in the mPFC. Scale bar, 50 µm. **b**, Segmentation of the neurons in **a**. **c**, Neuronal activity ($\Delta F/F$) extracted from the three example neurons outlined in **b**. **d**–**g**, Population activity traces (mean $\Delta F/F \pm$ s.e.m.) for neurons exhibiting decreased (**d**) or increased (**f**) activity during tone presentations in fear extinction session 3. Mean activity ($\Delta F/F$) during each task epoch (baseline, tone-on, tone-off; **e**) for the neuronal population depicted in **d** presents a significant decrease in activity (repeated-measures analysis of variance (ANOVA): main effect of time: $F(10,1600) = 3.138$, $P = 0.007$) but no significant difference between groups (group-by-time interaction: $F(10,1600) = 2.736$, $P = 0.1280$). NS, not significant (baseline, 0.285; tone-on, 0.595; tone-off, 0.578). Mean activity ($\Delta F/F$) during each task epoch (**g**) for the neuronal population depicted in **f** presents a significant increase in activity (repeated-measures ANOVA: main effect of time: $F(10,1770) = 4.945$, $P < 0.0001$) and a significant group-by-time interaction ($F(10,1770) = 3.806$, $P = 0.0008$). *$P = 0.013$, significant group difference in post-hoc contrast. NS, not significant (baseline, 0.128; tone-off, 0.601). Centre line, median; box, 25th and 75th percentiles; whiskers, minimum to maximum in **e**, **g**. **h**, Raster plot of neuronal activity for cells that encoded the timing of tones by increasing and decreasing activity in response to tone onset and offset, respectively. Each row represents one neuron. **i**, Population activity trace (mean $\Delta F/F \pm$ s.e.m.) for neurons depicted in **h**, timelocked to tone onset and averaged across tones. Repeated-measures ANOVA: main effect of time: $F(179,8234) = 7.033$, $P < 0.0001$; group by time interaction: $F(179,8234) = 2.749$, $P = 0.0093$. Data in **d**–**i** based on 1,204 total cells pooled from 3 independent experiments, from $n = 7$ control mice and $n = 8$ ABX mice.

To investigate whether the extinction learning deficits caused by altered microbiota are reversible, we recolonized previously germ-free (ex-GF) mice with a complete microbiota from healthy control mice at various developmental time points. Both ex-GF mice colonized when they were adults (ex-GF_adult) and ex-GF mice colonized at weaning age (ex-GF_weaning) still displayed impaired fear extinction compared to control mice (Fig. 5b, c), indicating that extinction learning deficits were not reversible in GF mice after weaning. However, when ex-GF mice were colonized immediately after birth by fostering to microbiota-replete specific-pathogen-free (SPF) surrogate mothers (ex-GF_fostered mice), their fear extinction behaviour was comparable to that of control_fostered mice (Fig. 5d), indicating that extinction learning and learning-related plasticity require microbiota-derived signals during a

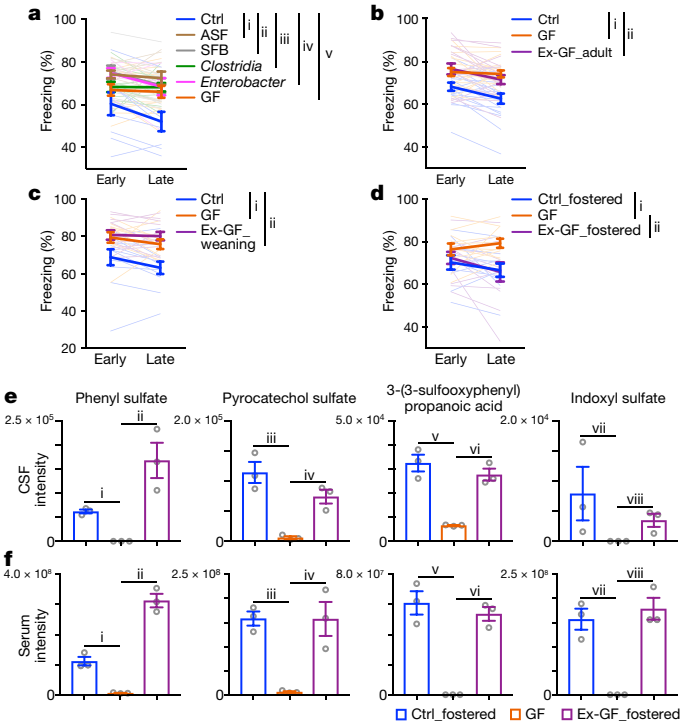



**Fig. 5 | Extinction learning deficits in GF mice are associated with alterations in microbiota-derived metabolites. a**–**d**, Fear extinction in control, GF and gnotobiotic mice colonized by SFB, *Clostridium*, *Enterobacter* or ASF bacteria (**a**); in control, GF and ex-GF_adult mice (**b**); in control, GF and ex-GF_weaning mice (**c**); or in control, GF and ex-GF_fostered mice (**d**) in the single-session 30-tone fear extinction assay. Data in **a**, **b**, **d** pooled from three independent experiments; data in **c** pooled from two independent experiments. **a**, $n = 9, 13, 9, 12, 7$ and 9 for control, GF, SFB, *Clostridium*, *Enterobacter* and ASF, respectively. **b**, $n = 18$ mice per group. **c**, $n = 12$ mice per group. **d**, $n = 10, 11$ and 12 for Ctrl_fostered, GF, and GF_fostered mice, respectively. Mean ± s.e.m.; AUC calculated for each mouse within each group followed by one-way ANOVA with Tukey's multiple comparisons test. **a**, $F(5,53) = 7.046$, $P = 4.10 \times 10^{-5}$. Adjusted $P$ values: i, $6.34 \times 10^{-5}$; ii, 0.0002; iii, 0.0042; iv, 0.0010; v, 0.0189. **b**, $F(2,51) = 11.92$, $P = 5.66 \times 10^{-5}$. Adjusted $P$ values: i, 0.0002; ii, 0.0005. **c**, $F(2,33) = 12.64$, $P = 8.40 \times 10^{-5}$. Adjusted $P$ values: i, 0.0016; ii, 0.0001. **d**, $F(2,30) = 5.131$, $P = 0.0121$. Adjusted $P$ values: i, 0.0228. ii, 0.0273. **e**, **f**, Relative abundances of four compounds in CSF (**e**) and serum (**f**) samples from Ctrl_fostered, GF and GF_fostered mice. $n = 3$ mice per group. Mean ± s.e.m.; unpaired two-sided $t$-tests were used. **e**, Adjusted $P$ values: i, $9.54 \times 10^{-5}$; ii, 0.0102; iii, 0.0036; iv, 0.0044; v, 0.0018; vi, 0.0011; vii, 0.1493; viii, 0.0331. **f**, Adjusted $P$ values: i, 0.0017; ii, 0.0002; iii, 0.0005; iv, 0.0128; v, 0.0014; vi, 0.0003; vii, 0.0020; viii, 0.0014.

critical developmental period before weaning. Lack of the microbiota in the neonatal period, no matter whether microbially colonized or not after weaning, renders deficits in fear extinction learning in adulthood (Fig. 5b–d). Whereas fear extinction was restored in ex-GF_fostered mice, we found no significant shift in the transcriptome of the mPFC of ex-GF_fostered mice or control_fostered mice (compared to GF mice) after the single-session fear extinction training (data not shown). This could reflect the shorter extinction model used in GF mice, which may induce smaller transcriptional changes than the protocol involving one session per day for three days that was used in ABX mice. Alternatively, the lack of transcriptional changes in the GF fostering studies could indicate the involvement of other processes such as post-translational or epigenomic modifications.

## Altered CSF metabolites and fear extinction

Next, we investigated whether changes in neuronal development and fear extinction learning were associated with alterations in microbiota-derived metabolites. We used untargeted comparative metabolomics of cerebrospinal fluid (CSF), serum and faecal samples from adult GF mice, control_fostered mice and ex-GF_fostered mice using high-resolution liquid chromatography and mass spectrometry (LC–MS). Using the xcms platform[40] for comparative analysis of the mass spectrometry datasets, we identified four metabolites—phenyl sulfate, pyrocatechol sulfate, 3-(3-sulfooxyphenyl)propanoic acid (all phenolic compounds) and indoxyl sulfate (Extended Data Fig. 10c)—that were significantly decreased in CSF, serum and faecal samples of GF mice compared to control_fostered mice, and were restored in ex-GF_fostered mice (Fig. 5e, f, Extended Data Fig. 10d). Downregulation of the same four metabolites was also detected when we compared GF with control CSF samples (Extended Data Fig. 10e), as well as in comparisons of ABX versus control samples (data not shown). Notably, 3-(3-hydroxyphenyl)-3-hydroxypropanoic acid (HPHPA, a derivative of 3-(3-sulfooxyphenyl) propanoic acid) and indoxyl sulfate have been associated with neuropsychiatric disorders such as impaired executive function, schizophrenia and autism in humans[41–43]. Other microbiota-derived phenolic compounds such as 4-ethylphenylsulfate have been reported to influence autism-related behaviours in mice[44], suggesting that these microbiota-derived compounds may have conserved roles in the development and function of neurons in multiple contexts.

## Discussion

This study informs a model in which alterations in exposure to the microbiota in neonatal and adult mice can have considerable and long-lasting effects on neuronal function and learning-related plasticity that subsequently regulate fear extinction behaviour (Extended Data Fig. 10f). From bulk RNA-seq and snRNA-seq data, the deficits in extinction learning correlate with malfunctions of the mPFC, notably in excitatory neurons. Transcranial two-photon live imaging confirmed the changes in neurons in the ABX mice (specifically, reduced extinction learning-associated spine formation and altered learning-related neuronal activity). Given that the vagus nerve does not contribute to the extinction learning deficits in ABX mice in this setting, the microbiota may affect the CNS through circulating microbiota-derived metabolites, directly influencing excitatory neurons in the mPFC and leading to deficits in extinction learning. In addition, microbiota-derived metabolites may also influence other cell subsets in the mPFC (such as microglia), and indirectly affect the excitatory neurons and behaviour. In accordance with this, and consistent with other studies[12], we found that the microglia in GF and ABX mice exhibited an immature state reminiscent of developing juvenile microglia, which may contribute to elevated spine pruning and reduced extinction-learning-associated spine formation. Future studies are necessary to determine the molecular basis of changes to cellular activity in the mPFC induced by alterations in microbiota-derived metabolites.

In summary, our findings offer one compelling explanation for the notable deficits in fear extinction learning in ABX and GF mice, and suggest that alterations in microbiota-derived metabolites contribute to altered neuronal activity and behaviour. Coupled with the growing literature on the influence of the microbiota, diet and lifestyle on chronic inflammatory diseases, metabolic health and cancer, the effects of beneficial microorganisms on brain health and behaviour highlight the need to better define the co-evolved relationship between the microbiota, the nervous system and mammalian behaviour.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1644-y.

1. Belkaid, Y. & Hand, T. W. Role of the microbiota in immunity and inflammation. *Cell* **157**, 121–141 (2014).
2. Hill, D. A. & Artis, D. Intestinal bacteria and the regulation of immune cell homeostasis. *Annu. Rev. Immunol.* **28**, 623–667 (2010).
3. Vuong, H. E., Yano, J. M., Fung, T. C. & Hsiao, E. Y. The microbiome and host behavior. *Annu. Rev. Neurosci.* **40**, 21–49 (2017).
4. Maren, S. Neurobiology of Pavlovian fear conditioning. *Annu. Rev. Neurosci.* **24**, 897–931 (2001).
5. VanElzakker, M. B., Dahlgren, M. K., Davis, F. C., Dubois, S. & Shin, L. M. From Pavlov to PTSD: the extinction of conditioned fear in rodents, humans, and anxiety disorders. *Neurobiol. Learn. Mem.* **113**, 3–18 (2014).
6. Mielcarz, D. W. & Kasper, L. H. The gut microbiome in multiple sclerosis. *Curr. Treat. Options Neurol.* **17**, 344 (2015).
7. Krajmalnik-Brown, R., Lozupone, C., Kang, D. W. & Adams, J. B. Gut bacteria in children with autism spectrum disorders: challenges and promise of studying how a complex community influences a complex disease. *Microb. Ecol. Health Dis.* **26**, 26914 (2015).
8. Zheng, P. et al. Gut microbiome remodeling induces depressive-like behaviors through a pathway mediated by the host's metabolism. *Mol. Psychiatry* **21**, 786–796 (2016).
9. Möhle, L. et al. Ly6C$^{hi}$ monocytes provide a link between antibiotic-induced changes in gut microbiota and adult hippocampal neurogenesis. *Cell Rep.* **15**, 1945–1956 (2016).
10. Hoban, A. E. et al. Regulation of prefrontal cortex myelination by the microbiota. *Transl. Psychiatry* **6**, e774 (2016).
11. Braniste, V. et al. The gut microbiota influences blood-brain barrier permeability in mice. *Sci. Transl. Med.* **6**, 263ra158 (2014).
12. Erny, D. et al. Host microbiota constantly control maturation and function of microglia in the CNS. *Nat. Neurosci.* **18**, 965–977 (2015).
13. Hoban, A. E. et al. The microbiome regulates amygdala-dependent fear recall. *Mol. Psychiatry* **23**, 1134–1144 (2018).
14. Lu, J. et al. Microbiota influence the development of the brain and behaviors in C57BL/6J mice. *PLoS ONE* **13**, e0201829 (2018).
15. Arentsen, T., Raith, H., Qian, Y., Forssberg, H. & Diaz Heijtz, R. Host microbiota modulates development of social preference in mice. *Microb. Ecol. Health Dis.* **26**, 29719 (2015).
16. Desbonnet, L., Clarke, G., Shanahan, F., Dinan, T. G. & Cryan, J. F. Microbiota is essential for social development in the mouse. *Mol. Psychiatry* **19**, 146–148 (2014).
17. Hepworth, M. R. et al. Group 3 innate lymphoid cells mediate intestinal selection of commensal bacteria-specific CD4$^{+}$ T cells. *Science* **348**, 1031–1035 (2015).
18. Pattwell, S. S. et al. Altered fear learning across development in both mouse and human. *Proc. Natl Acad. Sci. USA* **109**, 16318–16323 (2012).
19. Soliman, F. et al. A genetic variant BDNF polymorphism alters extinction learning in both mouse and human. *Science* **327**, 863–866 (2010).
20. Bravo, J. A. et al. Ingestion of *Lactobacillus* strain regulates emotional behavior and central GABA receptor expression in a mouse via the vagus nerve. *Proc. Natl Acad. Sci. USA* **108**, 16050–16055 (2011).
21. Sgritta, M. et al. Mechanisms underlying microbial-mediated changes in social behavior in mouse models of autism spectrum disorder. *Neuron* **101**, 246–259.e246 (2019).
22. Kipnis, J., Cohen, H., Cardon, M., Ziv, Y. & Schwartz, M. T cell deficiency leads to cognitive dysfunction: implications for therapeutic vaccination for schizophrenia and other psychiatric conditions. *Proc. Natl Acad. Sci. USA* **101**, 8180–8185 (2004).
23. Miyajima, M. et al. Metabolic shift induced by systemic activation of T cells in PD-1-deficient mice perturbs brain monoamines and emotional behavior. *Nat. Immunol.* **18**, 1342–1352 (2017).
24. Filiano, A. J. et al. Unexpected role of interferon-γ in regulating neuronal connectivity and social behaviour. *Nature* **535**, 425–429 (2016).
25. Tovote, P., Fadok, J. P. & Lüthi, A. Neuronal circuits for fear and anxiety. *Nat. Rev. Neurosci.* **16**, 317–331 (2015).
26. Dragunow, M. & Faull, R. The use of *c-fos* as a metabolic marker in neuronal pathway tracing. *J. Neurosci. Methods* **29**, 261–265 (1989).
27. Gale, G. D. et al. Role of the basolateral amygdala in the storage of fear memories across the adult lifetime of rats. *J. Neurosci.* **24**, 3810–3815 (2004).
28. Thion, M. S. et al. Microbiome influences prenatal and adult microglia in a sex-specific manner. *Cell* **172**, 500–516.e516 (2018).
29. Kierdorf, K. et al. Microglia emerge from erythromyeloid precursors via Pu.1- and Irf8-dependent pathways. *Nat. Neurosci.* **16**, 273–280 (2013).
30. Yuste, R. & Bonhoeffer, T. Morphological changes in dendritic spines associated with long-term synaptic plasticity. *Annu. Rev. Neurosci.* **24**, 1071–1089 (2001).
31. Yuste, R. & Denk, W. Dendritic spines as basic functional units of neuronal integration. *Nature* **375**, 682–684 (1995).
32. Trachtenberg, J. T. et al. Long-term *in vivo* imaging of experience-dependent synaptic plasticity in adult cortex. *Nature* **420**, 788–794 (2002).
33. Matsuzaki, M., Honkura, N., Ellis-Davies, G. C. & Kasai, H. Structural basis of long-term potentiation in single dendritic spines. *Nature* **429**, 761–766 (2004).
34. Lai, C. S., Franke, T. F. & Gan, W. B. Opposite effects of fear conditioning and extinction on dendritic spine remodelling. *Nature* **483**, 87–91 (2012).
35. Sudo, N. et al. Postnatal microbial colonization programs the hypothalamic-pituitary-adrenal system for stress response in mice. *J. Physiol. (Lond.)* **558**, 263–275 (2004).
36. Ivanov, I. I. et al. Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell* **139**, 485–498 (2009).
37. Atarashi, K. et al. Induction of colonic regulatory T cells by indigenous *Clostridium* species. *Science* **331**, 337–341 (2011).
38. Garrett, W. S. et al. Enterobacteriaceae act in concert with the gut microbiota to induce spontaneous and maternally transmitted colitis. *Cell Host Microbe* **8**, 292–300 (2010).
39. Schaedler, R. W. & Dubos, R. J. The fecal flora of various strains of mice. Its bearing on their susceptibility to endotoxin. *J. Exp. Med.* **115**, 1149–1160 (1962).
40. Tautenhahn, R., Patti, G. J., Rinehart, D. & Siuzdak, G. XCMS Online: a web-based platform to process untargeted metabolomic data. *Anal. Chem.* **84**, 5035–5039 (2012).
41. Shaw, W. Increased urinary excretion of a 3-(3-hydroxyphenyl)-3-hydroxypropionic acid (HPHPA), an abnormal phenylalanine metabolite of *Clostridia* spp. in the gastrointestinal tract, in urine samples from patients with autism and schizophrenia. *Nutr. Neurosci.* **13**, 135–143 (2010).
42. Keşli, R., Gökçen, C., Buluğ, U. & Terzi, Y. Investigation of the relation between anaerobic bacteria genus *Clostridium* and late-onset autism etiology in children. *J. Immunoassay Immunochem.* **35**, 101–109 (2014).
43. Yeh, Y. C. et al. Indoxyl sulfate, not p-cresyl sulfate, is associated with cognitive impairment in early-stage chronic kidney disease. *Neurotoxicology* **53**, 148–152 (2016).
44. Hsiao, E. Y. et al. Microbiota modulate behavioral and physiological abnormalities associated with neurodevelopmental disorders. *Cell* **155**, 1451–1463 (2013).




[1]Jill Roberts Institute for Research in Inflammatory Bowel Disease, Weill Cornell Medicine, Cornell University, New York, NY, USA. [2]Feil Family Brain and Mind Research Institute, Weill Cornell Medicine, Cornell University, New York, NY, USA. [3]Department of Psychiatry, Weill Cornell Medicine, Cornell University, New York, NY, USA. [4]Sackler Institute for Developmental Psychobiology, Weill Cornell Medicine, Cornell University, New York, NY, USA. [5]Department of Pharmacology, Weill Cornell Medicine, Cornell University, New York, NY, USA. [6]Boyce Thompson Institute and Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY, USA. [7]Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA, USA. [8]Center for Biomedical Science and Bioelectronic Medicine, Feinstein Institute for Medical Research, Northwell Health, Manhasset, NY, USA. [9]Elmezzi Graduate School, Feinstein Institute for Medical Research, Northwell Health, Manhasset, NY, USA. [10]Department of Surgery, Northshore University Hospital, Northwell Health, Manhasset, NY, USA. [11]Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Hempstead, NY, USA. [12]Howard Hughes Medical Institute, Koch Institute of Integrative Cancer Research, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA. [13]Friedman Center for Nutrition and Inflammation, Joan and Sanford I. Weill Department of Medicine, Department of Microbiology and Immunology, Weill Cornell Medicine, Cornell University, New York, NY, USA. *e-mail: col2004@med.cornell.edu; dartis@med.cornell.edu

# Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment, except for dendritic spine imaging data analysis, in which the raters were blinded to experimental conditions.

### Mice

C57BL/6J (Jax 664), *Rag1*$^{-/-}$ (Jax 2216), Thy1-YFP-H (Jax 3782) and BALB/c (Jax 651) mice were purchased from The Jackson Laboratory and bred in-house. Male mice were used at 7–16 weeks of age. In individual experiments, all mice were age-matched. All mice were maintained under SPF conditions on a 12-h light–dark cycle, and provided food and water ad libitum. Germ-free C57BL/6J mice and gnotobiotic mice were maintained at Weill Cornell Medical College, New York. All mouse experiments were approved by, and performed in accordance with, the Institutional Animal Care and Use Committee guidelines at Weill Cornell Medicine.

### Antibiotic treatment

Mice were provided autoclaved drinking water supplemented with a cocktail of broad-spectrum antibiotics as previously described[17]: ampicillin (0.5 mg/ml, Santa Cruz), gentamicin (0.5 mg/ml, Gemini Bio-Products), metronidazole (0.5 mg/ml Sigma), neomycin (0.5 mg/ml, Sigma), vancomycin (0.25 mg/ml, Chem-Impex International), and saccharin (4 mg/ml, Sweet'N Low, Cumberland Packing). Sweet'N Low was added to make the antibiotic cocktail more palatable. Antibiotic treatment was started two weeks before the experiments and continued for the duration of the experiments. Following antibiotic treatment mice exhibited no significant differences in weight gain, food or water intake (measured by Promethion metabolic cages) or perception of pain[45].

### Fear conditioning and extinction assays

Fear conditioning and extinction assays were performed as previously described[18,19]. For fear conditioning, mice were placed in shock chambers (Coulbourn Instruments), which were scented with 0.1% peppermint in 70% EtOH. After 2 min of habituation, mice were fear conditioned with 3 tone–shock pairings consisting of a 30-s (5 kHz, 70 dB) tone (conditioned stimulus, CS) that co-terminated with a 1-s (0.7 mA) foot shock (unconditioned stimulus, US). Intertrial intervals (ITIs) between each tone–shock pairing were 30 s. After the final tone–shock pairing, mice remained in the conditioning chambers for 1 min before being returned to their home cages.

For the classical 3-day (that is, 3-session) fear extinction, 24 h after fear conditioning, mice were placed in extinction chambers (different shape from the conditioning chambers), which were scented with 0.1% lemon in 70% EtOH. After 2 min of habituation, mice were exposed to 5 presentations of the tone (CS) in the absence of the shock (US). Each tone lasted for 30 s with an ITI of 30 s. After the final tone presentation, mice remained in the extinction chambers for 1 min before being returned to their home cages. Fear extinction sessions were repeated daily for three days.

For single-session 30-tone fear extinction, 20 min after fear conditioning, mice were placed in extinction chambers. After 2 min of habituation, mice were exposed to 30 presentations of the tone (CS) in the absence of the shock (US). Each tone lasted for 30 s with an ITI of 30 s. Extinction trials were binned into early and late sessions, with the early session representing the average of trials 1–15, and late trials representing the average of trials 16–30.

Experiments were controlled by Graphic State software (Coulbourn instruments). Mice were video-recorded for subsequent analysis.

### Fear behaviour

Mouse freezing behaviour was scored automatically using previously validated MATLAB code for automated phenotyping of mouse behaviour[46]. Per cent time spent freezing (freezing (%)) was calculated by dividing the amount of time spent freezing during the 30-s tone presentations by the duration of the tone.

### Immunofluorescence staining

Brain sections were prepared and stained for c-FOS, synaptophysin or PSD-95 expression as previously described[47]. All steps were carried out at room temperature unless otherwise specified. Ninety minutes after fear extinction session 3, mice were anaesthetized by intraperitoneal injection of Euthasol and perfused with PBS followed by 4% paraformaldehyde. Brains were collected, fixed in 4% paraformaldehyde overnight, and dehydrated in 30% sucrose at 4 °C. Coronal sections (40 μm) were cut using a sliding microtome frozen by powdered dry ice. Six sets of serial sections were collected in Eppendorf tubes each containing 2 ml cryoprotectant (30% glycerol and 30% ethylene glycol in 0.1 M sodium phosphate, pH 7.4) and stored at −20 °C. Free-floating serial sections (1 in every 3) were washed in TBS, incubated for 30 min in blocking buffer (4% normal horse serum, 1% BSA and 0.2% Triton X-100 in TBS) and incubated overnight at 4 °C with rabbit anti-c-FOS primary antibody (sc-52, Santa Cruz), or mouse anti-synaptophysin (SAB4200544, Sigma-Aldrich) or PSD-95 (7E3-1B8, Sigma-Aldrich) diluted 1:1,000 in the blocking buffer. Sections were then washed in TBS and incubated for 2 h with Alexa-Fluor-555-labelled donkey anti-rabbit or anti-mouse antibody (Invitrogen) diluted 1:500 in TBS with 0.2% Triton X-100. Sections were again washed, mounted on chromalum/gelatin-coated slides and air-dried for 2 h in darkness. Slides were coverslipped with water-soluble glycerol-based mounting medium containing DAPI and sealed with nail polish.

Estimation of cell densities of c-FOS$^{+}$ neurons in BLA and IL was performed with StereoInvestigator 9.0 (MicroBrightField). In brief, serial sections (every third section, 120 μm) were numbered by rostra-caudal order, and contours of BLA and IL were traced by referring to the Allen Brain Atlas (Allen Institute). All cells across all sections per mouse were counted. Individual cell density was calculated for each mouse by dividing the total sampled cell numbers by the total volume of the region.

For synaptophysin and PSD-95 images, confocal microscopy was performed with a Zeiss LSM 880 Laser Scanning Confocal Microscope using 63× oil immersion lens. Images were acquired with 2× digital zoom. Image stacks were 5 μm in thickness with *z*-step size of 0.5 μm, and were analysed using ImageJ software (http://rsbweb.nih.gov/ij/).

### Intracranial window implantation

Mice were anaesthetized with isoflurane (induction, 5%; maintenance, 1–2%) and administered dexamethasone (1 mg/kg, i.p.) to reduce brain swelling and metacam (2 mg/kg, i.p.) as a prophylactic analgesic. Scalp fur was trimmed, and the skull surface was exposed with a midline scalp incision. Bupivicaine (0.05 ml, 5 mg/ml) was administered topically as a second prophylactic analgesic. A circular titanium head plate was positioned over the region to be imaged (1.7 mm anterior to the bregma suture and centred over the midline) using dental cement (Metabond). A high-speed dental drill (Model EXL-M40, Osada) and 0.5-mm burr were used to open a small (about 4-mm) craniotomy. A 3-mm round coverslip (Warner Instruments) was lowered through the craniotomy to rest on top of the brain using a digital micromanipulator. The window was then fixed to the skull using veterinary adhesives (first Vetbond, then Metabond).

### Viral injection

AAV5–hSyn–GCaMP6s was obtained from the UPenn Vector Core. Viral injection surgeries were performed with mice (8–10 weeks of age) under isoflurane anaesthesia (induction, 5%; maintenance, 1–2%) with regular monitoring for stable respiratory rate and absent tail pinch response. The scalp was shaved, and mice were fixed in a stereotactic frame (Kopf Instruments) with non-rupturing ear bars. A heating pad was used to prevent hypothermia. A midline incision was made to expose the skull and bupivacaine was applied onto the skull for local anaesthesia. Virus injections (1,000 nl) were delivered with a 10-μl Hamilton syringe and

33-gauge bevelled needle, injected at 100 nl/min using an injection pump (World Precision Instruments). Injection coordinates relative to Bregma were: 1.7 mm anterior, 0.4 mm lateral, and 1.3 mm ventral. Following injection, the injection needle was held at the injection site for 2 min then slowly withdrawn. The skin was then closed using Vetbond (3M) and the mice recovered on a heating pad before being returned to their home cages.

### Transcranial two-photon imaging

Dendritic spine imaging was conducted as previously described[48]. In brief, image stacks of dendritic segments were acquired using a two-photon laser scanning microscope (Olympus RS) equipped with a scanning galvanometer and a Spectra-Physics Mai Tai DeepSee laser tuned to 920 nm, and a 25×, long working distance water-immersion microscope objective (NA = 1.05, Olympus). Fluorescence was detected through gallium arsenide phosphide photomultiplier tubes using the Fluoview acquisition software (Olympus), and images were collected in the green channel using an F30FGR bandpass filter (Semrock). All imaging experiments began by obtaining a low-magnification *z*-stack (no digital zoom) to aid in relocating the same sites repeatedly over time, in conjunction with vascular landmarks and the contours of the prism. For spine imaging experiments, we acquired *z*-stacks (512 × 512 pixels, 2-µs pixel dwell time, 0.75–1-µm step size) with 3× digital zoom through up to 250 µm of tissue in *z*. Spine imaging experiments occurred under KX anaesthesia (ketamine 100 mg/ml and xylazine 10 mg/ml, at dosages of 0.1 ml/10 g body weight). For calcium imaging experiments, we acquired time-lapse images (512 × 512 pixels, 3 frames per second, about 1,450 frames,) spanning an area of the mPFC measuring approximately 508 µm by 508 µm. All calcium imaging experiments were carried out on awake mice. For repeated imaging over intervals of days, the procedure above was repeated, and the region to be imaged was identified by referring to vascular landmarks and the contours of the cranial window.

### Spine imaging analysis

Spine remodelling dynamics were quantified as previously described[48]. Image stacks were analysed using ImageJ software (http://rsbweb.nih.gov/ij/). Raters blinded to experimental condition compared pairs of images of the same dendritic segment and identified stable spines (present in image 1 and 2), eliminated spines (present in image 1 but not in image 2) and formed spines (present in image 2 but not in image 1), each quantified as a percentage of the total number of spines identified in the initial image. Filopodia were defined as dendritic protrusions with a length exceeding three times their maximum width and were excluded from spine remodelling analyses.

### Calcium imaging analysis

**Preprocessing.** We used standard, validated procedures for preprocessing and analysing calcium imaging time series data. *X*–*Y* motion artefacts were corrected using the ImageJ plugin 'Image Stabilizer' created by K. Li and S. Kang (https://www.cs.cmu.edu/~kangli/code/Image_Stabilizer.html). Image time series were segmented into individual cells using custom MATLAB scripts based on an established sorting algorithm that combines independent components analysis and image segmentation based on threshold intensity, variance, and skewness in the *X*–*Y* motion-corrected dataset[49–51]. Image segmentation results were manually inspected for quality control. Fluorescence signal time series ($\Delta F/F$: change in fluorescence divided by baseline fluorescence) were calculated for each individual neuronal segment: a 40-s sliding window was used to calculate the baseline fluorescence for each cell, accounting for both differences in GCaMP expression and de-trending for slow timescale changes in fluorescence[50].

**Analysis.** First, we tested for cells exhibiting tone-sensitive activity, using repeated-measures ANOVA to identify cells with a statistically significant increase (Fig. 4f) or decrease (Fig. 4d) in activity during tone presentations, compared to their activity during a two-minute pre-tone baseline period. To estimate statistical significance while accounting for autocorrelation in calcium transient time series and correcting for multiple comparisons, we repeated this analysis 10,000 times for each cell after shuffling the timing of the baseline period and the timing of the tone onsets and selected a *P* value threshold to limit the FDR to less than 5%. Next, to test for group (ABX versus control) effects on activity in each of these cell populations (Fig. 4e, g), we used a two-factor (time and group) repeated-measures ANOVA and post-hoc linear contrasts to test for between-group differences in activity during each task epoch (baseline, tone on and tone off). Finally, to test for cells that also encoded the precise timing of the tones, exhibiting tone-locked activity that increased or decreased in response to the onset of each tone, we used a procedure analogous to the one described above, using repeated-measures ANOVA to test for changes in activity in the tone on versus tone off epochs; estimating statistical significance in shuffled data as above; and testing for group effects on activity using a two-factor (time and group) repeated-measures ANOVA (Fig. 4h, i).

### RNA-seq

Mouse mPFC was dissected by referring to the Allen Brain Atlas. Coordinates relative to bregma are: 1.3 to 2.8 mm anterior, −1 to 1 mm lateral, and 0 to 1 mm ventral. RNA was extracted using Trizol (Invitrogen) and chloroform and further purified using the RNAeasy mini spin columns (Qiagen). RNA-seq libraries were prepared and sequenced by the Epigenomics Core at Weill Cornell Medicine on an Illumina HiSeq 2500, producing 50-bp single-end reads. Sequenced reads were demultiplexed using CASAVA v1.8.2 and adapters trimmed using FLEXBAR v2.4[52].

### RNA-seq analysis

Sequenced reads were aligned to the mouse genome GRCm38/mm10 using STAR v2.3.0[53]. Reads counts at the gene level were calculated using Rsubread[54]. Normalization for library size and differential expression analysis were performed using DESeq2[55] v1.18. Only genes with at least ten raw reads in each sample were tested for differential expression. PERMANOVA[56] was used to test whether antibiotic treatment accounted for a significant portion of the variance in gene expression after fear extinction. Specifically, expression of the 500 genes with the highest variance (after applying the variance stabilization transformation of DESeq2[55]) was analysed using the adonis function of the vegan R package (https://CRAN.R-project.org/package=vegan) using the Euclidean metric and 20,000 permutations. Differentially expressed genes were used for GO enrichment analysis (http://www.pantherdb.org/)[57], KEGG analysis (https://www.genome.jp/kegg)[58] and STRING analysis (https://string-db.org/)[59].

### Brain-resident immune cell isolation and flow cytometry

Brain-resident immune cells were isolated using Percoll gradients[60]. Mice were anaesthetized and perfused with ice-cold HBSS. Brains were removed, homogenized, resuspended with 30% Percoll, and layered on top of 70% Percoll. After centrifugation (500*g*, 30 min), immune cells gathered in the 30–70% interphase.

For flow cytometric analyses, cells were washed, incubated with purified anti-mouse CD16/CD32 (clone 93, Biolegend) to block the Fc receptors, and then stained with anti-CD45 (clone 30-F11, Biolegend), anti-CD4 (clone RM4-5, eBioscience), anti-CD8a (clone 53-6.7, BD Biosciences), anti-CD19 (clone 1D3, eBioscience), anti-CD11b (clone M1/70, eBioscience), anti-CD11c (clone N418, eBioscience), anti-F4/80 (clone BM8, eBioscience), anti-LY6G (clone 1A8-Ly6g, eBioscience), anti-LY6C (clone HK1.4, eBioscience) and anti-CSF1R (clone AFS98, eBioscience). Data were collected on a LSRFortessa cytometer (BD Biosciences) and analysed with FlowJo software (Tree Star). Dead cells were excluded from analyses based on LIVE/DEAD Fixable Aqua dead cell staining (Invitrogen). Non-singlet events were excluded from analyses based on the side scatter height (SSC-H) versus side scatter width (SSC-W), and

then the forward scatter height (FSC-H) versus forward scatter width (FSC-W) characteristics.

### Microbial colonization

For ex-GF mice colonized when they were adults (ex-GF_adult), dirty bedding from SPF mice was placed in the GF cages of 8-week-old GF mice two weeks before the fear conditioning and extinction assay. For ex-GF mice colonized when they were weaned (ex-GF_weaning), 3-week-old GF mice were co-housed with 3-week-old SPF mice until they were 8 weeks old and then subjected to the fear conditioning and extinction assay.

### Fostering of pups

C57BL/6J GF and control SPF newborn pups were fostered by BALB/c mothers until weaning.

### 16S qPCR

DNA was isolated from faecal samples of control and ABX mice using the DNeasy PowerSoil kit (Qiagen). Equal amounts of purified faecal DNA (4 ng per reaction) were added to qPCR reactions with universal 16S primers using SYBR green chemistry (UniF340: 5′-ACTCCTACGG-GAGGCAGCAGT-3′; UniR514: 5′-ATTACCGCGGCTGCTGGC-3′). 16S DNA levels in each sample were normalized to the average of the control mouse group.

### 16S amplicon sequencing and analysis

16S rRNA gene sequencing methods were adapted from the methods developed for the NIH-Human Microbiome Project (https://hmpdacc.org/). In brief, bacterial genomic DNA was extracted using MO BIO PowerSoil DNA Isolation Kit (MO BIO Laboratories). The 16S rDNA V4 region was amplified by PCR and sequenced in the MiSeq platform (Illumina) using the 2 × 250-bp paired-end protocol. Raw reads were processed and clustered into operational taxonomic units (OTUs) using USEARCH version 11[61]. Specifically, reads were demultiplexed and read pairs merged, with a maximum of five mismatching bases in the overlap region, as well as a minimum sequence agreement of 80%. PhiX contaminant sequences were removed, and merged sequences were filtered according to FASTQ quality scores using a maximum expected error number of 0.1. Filtered sequences were clustered into OTUs at a 97% identity threshold using the USEARCH cluster_otus command with default settings. Merged reads (unfiltered) were mapped to the OTU representative sequences, generating an OTU table. Taxonomic classification of OTU representative sequences was performed with the USEARCH SINTAX command with a confidence score of 0.8, using version 16 of the RDP 16S training set[62]. Diversity estimation and principal coordinates analysis (PCoA) ordination were performed using the phyloseq R package[63] after subsampling the OTU table to even depth.

### snRNA-seq

Nuclei were extracted from four frozen mPFC samples (two from ABX mice and two from control mice) with a glass dounce tissue grinder set (Millipore Sigma no. D8938) and Nuclei EZ Prep (Millipore Sigma no. NUC101-1kt). Each sample was dounced with pestles A and B (24× each) in 2 ml EZ prep buffer, washed with 5 ml EZ prep, and resuspended in 1 ml resuspension buffer (1× PBS, 0.1% BSA, 25 U/ml recombinant RNase inhibitor, Takara 2313B). Single nucleus suspensions were strained through a 35-μm cell strainer (Corning 352235), visually inspected under a microscope, and loaded onto 3′ library chips as per the manufacturer's protocol for the Chromium Single Cell 3′ Library & Gel Bead kit (v.3) (10X Genomics 1000092). For each sample, an input of 11,000 nuclei was added to each channel. Libraries were sequenced at a mean depth of 21,714 reads per nucleus on a HiSeqX.

### snRNA-seq data processing

Demultiplexed FASTQ files were generated using Cell Ranger v2.0. Reads were aligned to the mm10 mouse transcriptome containing pre-mRNA annotations, similarly to previously described[64], to generate raw gene expression matrices (nuclei by genes). Expression matrices across all four samples were merged and loaded into Scanpy (version 1.4.0)[65]. Genes found in fewer than three nuclei were filtered out. Nuclei were filtered out using the following criteria: fewer than 600 genes (likely to be empty droplets), more than 5,000 genes (likely to be doublets), >2% of reads mapping to mitochondrial genes, >0.1% of reads mapping to caspase genes to remove apoptotic cells. The resulting filtered matrix consisted of 38,649 nuclei and 22,451 genes. The filtered gene expression matrix was normalized within each nucleus, resulting in a filtered, nuclei-normalized matrix $X$, then log-normalized by calculating $\ln(X+1)$. Before selecting variable genes, we masked genes that contain highly repetitive regions in intronic regions that result in inflated read counts (PISD, Mylip, Gm17660) and highly expressed lncRNAs that affect within-nuclei normalization (Gm28928, Malat1). We selected 1,535 highly variable genes using the highly_variable_genes module in scanpy (min_mean = 0.1, max_mean = 3, min_disp = 0.8) for clustering analysis.

### snRNA-seq data clustering

We first regressed out the number of UMIs and the number of genes. Each gene was then scaled to unit variance. We then conducted dimensionality reduction via PCA using the ARPACK SVD solver in scanpy, computed the $k$-nearest neighbour graph with PCs 1–40 and $k = 30$ nearest neighbours. Clusters were determined with unsupervised clustering using the Louvain algorithm[66,67] and resulted in 24 clusters. Differential expression analysis was conducted to find the top 100 genes enriched in each cluster with the rank_genes_groups module in scanpy using logistic regression[68]. We annotated clusters post hoc on the basis of known marker genes[69,70] among the top 100 enriched genes. For visualization, we embedded the profiles with UMAP (uniform manifold approximation and projection[71]).

### Cell-type-specific differential expression analysis

To find DEGs between ABX and control mice for each cluster, we used statsmodels in Python to implement a mixed linear model for each cluster $c$. Specifically, we used the regression $Y_{i,c} \sim T + N + (1|B)$, in which $Y_{i,c}$ is the $\ln(X+1)$ expression vector for gene $i$ across all nuclei in cluster $c$, $T$ is a binary variable reflecting membership of the nucleus in either ABX or control sample, $N$ is the number of genes detected in each nucleus, and $B$ is a categorical variable denoting the 10× channel used for each sample to control for batch effects. We used a Bonferroni-corrected $P$ value of $10^{-7}$ as the cut-off for significance. For plotting, we used DEGs that had a minimum $\log_2$(fold change) of 0.31 (absolute fold change of 1.24) in either direction, and independently found to be significant in at least two clusters. To rank clusters based on the number of DEGs (Extended Data Fig. 5b), we first randomly sampled 500 nuclei with replacement from each cluster to maintain comparable statistical power across clusters and re-ran the mixed linear model as described; all other plots and analyses, including GO enrichment, were based on the full list of DEGs obtained without downsampling.

### Vagotomy

The following subdiaphragmatic vagotomies and pyloromyotomy procedure were modified from previously described procedures[72]. Mice were anaesthetized via IP injection of a ketamine (144 mg/kg)/xylazine (13 mg/kg) cocktail. A midline incision was made and the stomach was retracted inferiorly to expose the distal oesophagus and the gastroesophageal junction. The anterior (left) and posterior (right) branches of the vagus nerve were identified running alongside the oesophagus and severed distal to the hepatic branches. The stomach was then placed back into the anatomical position and a pyloromyotomy was performed using a bent 23-gauge needle. The superficial muscular layers were incised in a longitudinal fashion and closed transversely with 4-0 vicryl sutures. The peritoneum was then closed with a running 4-0 vicryl suture and the skin approximated with staples. Mice were allowed to recover from anaesthesia under a heat lamp and returned to the colony room

once awake and ambulating. For non-vagotomized mice, the vagus nerve was gently exposed without further manipulation. Mice were monitored for seven days. The completeness of subdiaphragmatic vagotomy was verified by examining fluorescent label of the dorsal motor vagal nucleus (DMV) on brainstem sections one week after intraperitoneal injection of FluoroGold. The absence of fluorescent label in DMV neurons was accepted as a marker of complete vagotomy.

### Mass spectrometry

High-resolution LC–MS analysis was performed on a Dionex 3000 UPLC coupled with a Thermo Q-exactive high-resolution mass spectrometer equipped with a HESI ion source. Metabolites were separated using a water–acetonitrile gradient on a Agilent Zorbax Eclipse XDB-C18 column (150 mm × 2.1 mm, particle size 1.8 µm) maintained at 40 °C; solvent A: 0.1% formic acid in water; solvent B: 0.1% formic acid in acetonitrile. The A–B gradient started at 1% B for 1 min after injection and increased linearly to 100% B at 15 min, using a flow rate of 0.5 ml/min. Mass spectrometer parameters: spray voltage 2.9 kV, capillary temperature 320 °C, prober heater temperature 300 °C; sheath, auxiliary, and spare gas 70, 2, and 0 ml/min, respectively; S-lens RF level 55, resolution 140,000 at $m/z$ 200, AGC target $1 \times 10^6$. The instrument was calibrated weekly with positive and negative ion calibration solutions (Thermo-Fisher). Each sample was analysed in negative and positive modes using a $m/z$ range of 100 to 1,500.

### Feature detection, characterization and compound synthesis

LC–MS RAW files from triplicate faecal, serum and CSF samples from adult ex-GF_fostered, Ctrl_fostered and GF mice were converted to mzXML (profile mode) using MSConvert (ProteoWizard), followed by analysis using a customized XCMS R-script based on the centWave XCMS algorithm to extract features[40]. Resulting tables of all detected features were used to compute ex-GF_fostered mice versus GF mice and Ctrl_fostered mice versus GF mice peak area ratios. To select differential features, we applied a filter that retained entries with peak area ratios larger than 2 (down in GF mice) or smaller than 0.5 (up in GF mice). We manually curated the resulting list to remove false positive entries—that is, features that upon manual inspection of raw data were not differential. For the features that were verified to be differential, we examined elution profiles, isotope patterns, and MS1 spectra to find molecular ions and remove adducts, fragments, and isotope peaks.

The structures of the four differential compounds were confirmed by coinjection with synthesized or commercial samples. Phenyl sulfate and indoxyl sulfate were purchased from TCI America and Sigma-Aldrich, respectively. Pyrocatechol sulfate and 3-(3-sulfooxyphenyl)propanoic acid were prepared following a previously published procedure[73]. To a stirred solution of catechol (Sigma-Aldrich, 0.55 g, 5 mmol) or 3-(4-hydroxyphenyl)propionic acid (Sigma-Aldrich, 0.88 g 5 mmol) in dry pyridine (2.5 ml), sulfur trioxide pyridine complex (0.88 g, 6 mmol) was added at room temperature. The resulting mixtures were heated in an oil bath at 45 °C and stirred for 2 h. The reactions were then allowed to cool to ambient temperature and transferred separately to flasks each containing 25 ml of 1 N KOH cooled in an ice bath. To each of the aqueous mixtures was added 100 ml of 2-propanol and the two reactions were left at 4 °C for 16 h. At this point, the products were filtered off as white precipitates. The crude products were taken up in 50 ml (3:1 ethanol:water) and heated to reflux, hot filtered, and placed in the fridge for recrystallization. This last step was then repeated. Totals of 210 and 270 mg of pyrocatechol sulfate and 3-(3-sulfooxyphenyl)propanoic acid were obtained, corresponding to yields of about 20%.

### Statistical analysis

Statistical tests were performed with Prism (GraphPad). Unless specifically indicated otherwise, Student's $t$-tests were used to compare endpoint means of different groups. Error bars depict the s.e.m.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Code availability

The algorithm used for automated scoring of freezing behaviour is available at https://www.seas.upenn.edu/~molneuro/software.html. The algorithm used for motion artefact correction in 2P calcium imaging data is available at http://www.cs.cmu.edu/~kangli/code/Image_Stabilizer.html. All other analysis code is available from the corresponding author upon reasonable request.

## Data availability

RNA-seq data, 16S rRNA-seq data and snRNA-seq data are available at Gene Expression Omnibus and BioProject under accession numbers GSE134808, PRJNA556230 and GSE135326, respectively. All datasets generated and/or analysed during the current study are presented in this published article, the accompanying Source Data or Supplementary Information, or are available from the corresponding author upon reasonable request.


45. Kang, M. et al. The effect of gut microbiome on tolerance to morphine mediated antinociception in mice. *Sci. Rep*. **7**, 42658 (2017).
46. Patel, T. P. et al. An open-source toolbox for automated phenotyping of mice in behavioral tasks. *Front. Behav. Neurosci*. **8**, 349 (2014).
47. Dincheva, I. et al. FAAH genetic variation enhances fronto-amygdala function in mouse and human. *Nat. Commun*. **6**, 6395 (2015).
48. Liston, C. et al. Circadian glucocorticoid oscillations promote learning-dependent synapse formation and maintenance. *Nat. Neurosci*. **16**, 698–705 (2013).
49. Mukamel, E. A., Nimmerjahn, A. & Schnitzer, M. J. Automated analysis of cellular signals from large-scale calcium imaging data. *Neuron* **63**, 747–760 (2009).
50. Dombeck, D. A., Khabbaz, A. N., Collman, F., Adelman, T. L. & Tank, D. W. Imaging large-scale neural activity with cellular resolution in awake, mobile mice. *Neuron* **56**, 43–57 (2007).
51. Rajasethupathy, P. et al. Projections from neocortex mediate top-down control of memory retrieval. *Nature* **526**, 653–659 (2015).
52. Dodt, M., Roehr, J. T., Ahmed, R. & Dieterich, C. FLEXBAR—flexible barcode and adapter processing for next-generation sequencing platforms. *Biology (Basel)* **1**, 895–905 (2012).
53. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
54. Liao, Y., Smyth, G. K. & Shi, W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res*. **41**, e108 (2013).
55. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. **15**, 550 (2014).
56. Anderson, M. J. A new method for non-parametric multivariate analysis of variance. *Austral Ecol*. **26**, 32–46 (2001).
57. Mi, H., Muruganujan, A., Casagrande, J. T. & Thomas, P. D. Large-scale gene function analysis with the PANTHER classification system. *Nat. Protocols* **8**, 1551–1566 (2013).
58. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. **28**, 27–30 (2000).
59. Snel, B., Lehmann, G., Bork, P. & Huynen, M. A. STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. *Nucleic Acids Res*. **28**, 3442–3444 (2000).
60. Pino, P. A. & Cardona, A. E. Isolation of brain and spinal cord mononuclear cells using percoll gradients. *J. Vis. Exp*. **48**, 2348 (2011).
61. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
62. Cole, J. R. et al. Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res*. **42**, D633–D642 (2014).
63. McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**, e61217 (2013).
64. Bakken, T. E. et al. Single-nucleus and single-cell transcriptomes compared in matched cortical cell types. *PLoS ONE* **13**, e0209648 (2018).
65. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol*. **19**, 15 (2018).
66. Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech*. **2008**, P10008 (2008).
67. Traag, V. A. Faster unfolding of communities: speeding up the Louvain algorithm. *Phys. Rev. E* **92**, 032801 (2015).
68. Ntranos, V., Yi, L., Melsted, P. & Pachter, L. A discriminative learning approach to differential expression analysis for single-cell RNA-seq. *Nat. Methods* **16**, 163–166 (2019).
69. Zhang, Y. et al. An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J. Neurosci*. **34**, 11929–11947 (2014).
70. Habib, N. et al. Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat. Methods* **14**, 955–958 (2017).

71. McInnes, L., Healy, J., Saul, N. & Großberger, L. UMAP: uniform manifold approximation and projection. *J. Open Source Softw.* **3**, 861 (2018).
72. Dezfuli, G. et al. Subdiaphragmatic vagotomy with pyloroplasty ameliorates the obesity caused by genetic deletion of the melanocortin 4 receptor in the mouse. *Front. Neurosci.* **12**, 104 (2018).
73. Edwards, D. R., Lohman, D. C. & Wolfenden, R. Catalytic proficiency: the extreme case of S-O cleaving sulfatases. *J. Am. Chem. Soc.* **134**, 525–531 (2012).

**Acknowledgements** We thank the members of the Artis and Liston laboratories for discussion and reading of the manuscript; G. Eraslan, S. Simmons, and C. Smillie for discussions about the snRNA-seq analyses; and the Metabolic Phenotyping Center of Weill Cornell Medicine for technical advice and support. The summary cartoon (Extended Data Fig. 10f) was created with BioRender.com. This work was supported by the Jill Roberts Institute (to G.G.P.), JSPS Overseas Research Fellowships (to S.M.), the National Institute of General Medical Sciences (1R35GM118182-01, to K.J.T.), the National Institute of Allergy and Infectious Diseases (1P01AI102852-01A1, to K.J.T. and S.S.C.), the National Institutes of Health (NS052819, to F.S.L), the Rita Allen Foundation, the One Mind Institute, the Klingenstein-Simons Foundations, the Brain and Behavior Research Foundation, and the National Institutes of Mental Health (R01 MH109685, R01 MH118451) (all to C.L.), the National Institutes of Health (AI074878, AI095466, AI095608 and AI102942), the Burroughs Wellcome Fund, the Crohn's and Colitis Foundation and the Rosanne H. Silbermann Foundation (all to D.A.). F.C.S. is a Faculty Scholar of the Howard Hughes Medical Institute. This work was supported by the Klarman Cell Observatory at the Broad Institute. A.R. is a Howard Hughes Medical Institute Investigator.

**Author contributions** C.C. carried out most of the experiments and analysed the data. M.H.M., D.J., T.H.W., H.C., A.M.K., T.T., M.E.A., L.Z., N.J.B., R.Y., S.M., C.N.P., A.L., H.C.M., F.T., S.S.C., K.J.T., A.R., F.C.S. and F.S.L. helped with experiments. H.C. and A.R. performed snRNA-seq and analysis. G.G.P. performed bulk RNA-seq and 16S rDNA-seq analysis. D.A., C.L. and C.C. conceived the project, analysed data, and wrote the manuscript with input from all co-authors.

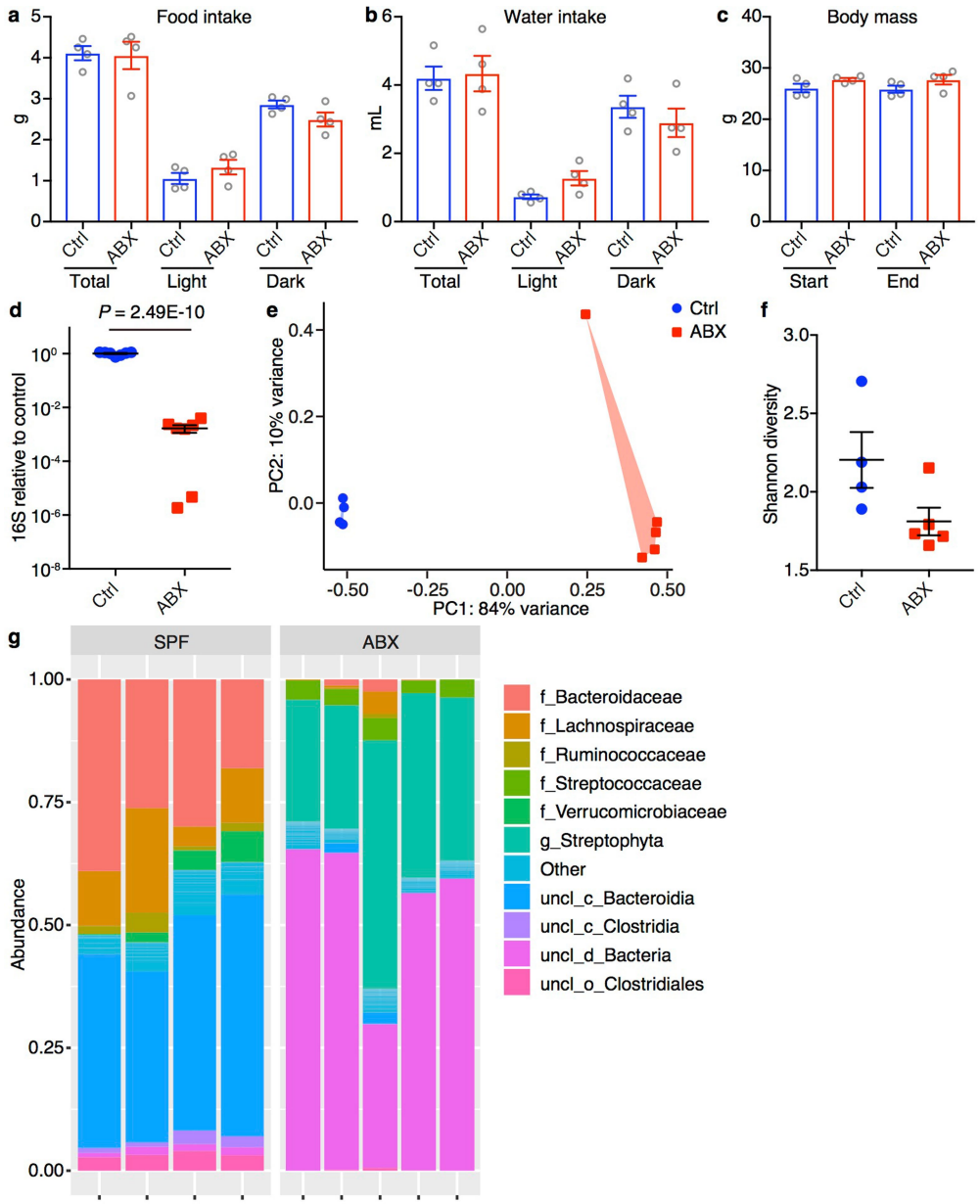
**Competing interests** A.R. is an SAB member of ThermoFisher Scientific and Syros Pharmaceuticals and a co-founder and equity holder of Celsius Therapeutics. D.A. has contributed to scientific advisory boards at MedImmune, Pfizer, FARE, and the KRF. The other authors declare no competing interests.

**Additional information**

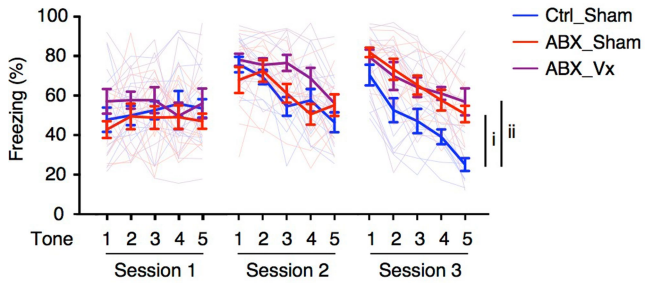
**Supplementary information** is available for this paper at https://doi.org/10.1038/s41586-019-1644-y.

**Correspondence and requests for materials** should be addressed to C.L. or D.A.

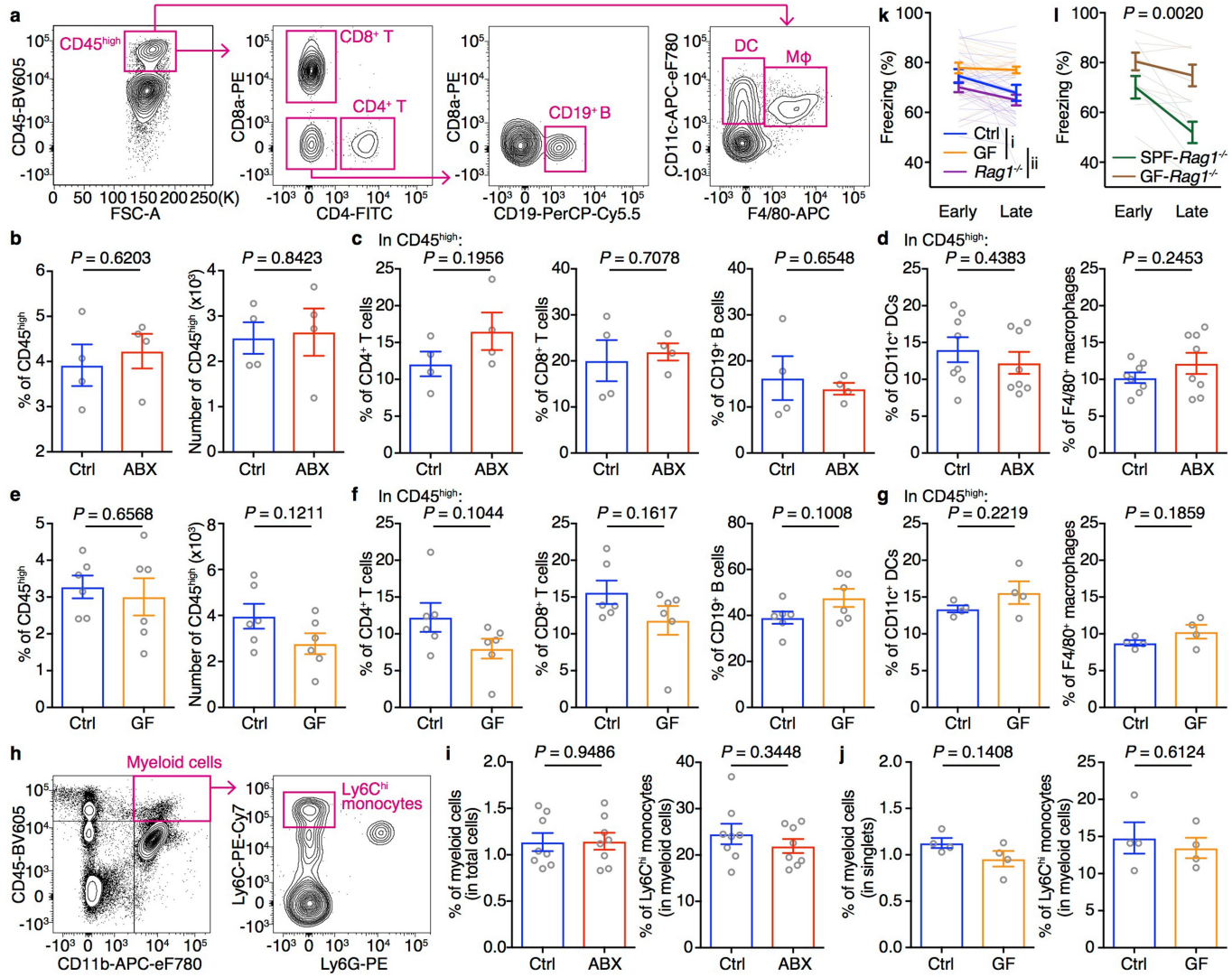
**Peer review information** *Nature* thanks Drew Kiraly and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



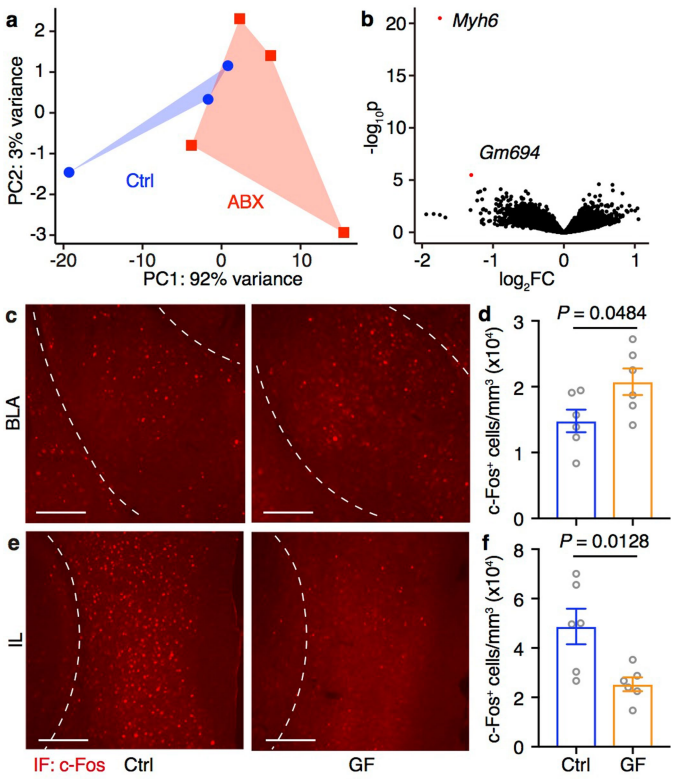

**Extended Data Fig. 1 | Antibiotic treatment results in bacterial community restructuring. a**–**c**, Food intake (**a**), water intake (**b**) and weight gain (**c**) of the mice measured using the Promethion Metabolic Cage System. Antibiotic treatment was started two weeks before the experiment and continued for the duration of the experiment. For food (**a**) and water intake (**b**), the mice were acclimated to the system for the first four days followed by one day of data collection. Body mass (**c**) of the mice was measured at the beginning (Start) and the end (End) of the 5-day experiment. $n = 4$ mice per group. Mean ± s.e.m. Total, full day. Light and Dark denote the light and dark periods of the 12-h cycle. **d**, 16S rDNA gene copies as quantified by real-time PCR with reverse transcription (RT–PCR) from stool pellets collected from control or ABX mice. Data pooled from two independent experiments. $n = 7$ mice per group. Mean ± s.e.m.; unpaired two-sided $t$-test. **e**–**g**, PCoA (**e**), alpha-diversity Shannon index (**f**) and taxonomic classification (**g**) of 16S rDNA in stool pellets collected from control or ABX mice. Control $n = 4$, ABX $n = 5$. For PCoA plot PERMANOVA: $F = 33.579$, $Df = 1$, $P = 0.00804$. For phylogenetic classification 'f_', 'g_', 'uncl_c_', 'uncl_d_' and 'uncl_o_' stand for 'family_', 'genus_', 'unclassified_class_', 'unclassified_domain_' and 'unclassified_order_', respectively. 'uncl_d_Bacteria' matches exactly to mitochondria or chloroplasts, probably from the food. Mean ± s.e.m. in **f**.



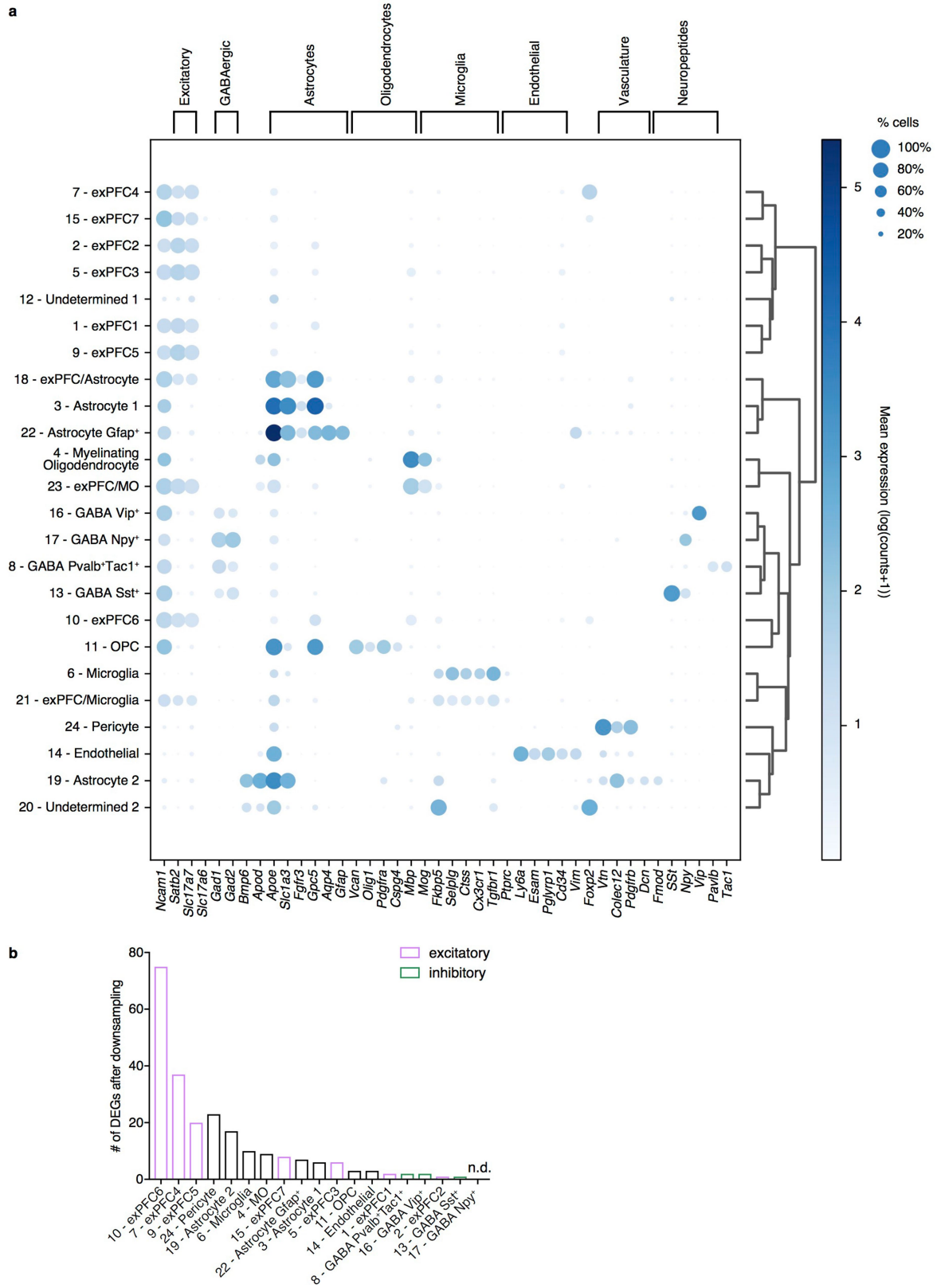

**Extended Data Fig. 2 | Antibiotic-treated mice retain deficits in extinction learning after vagotomy.** Fear extinction in sham-operated control (Ctrl_Sham) or ABX mice (ABX_Sham) and in vagotomised ABX mice (ABX_Vx) mice over the course of 3 days or sessions. Ctrl_Sham $n = 10$, ABX_Sham $n = 10$, ABX_Vx $n = 12$. Mean ± s.e.m.; AUC was calculated for each mouse within each group, followed by unpaired two-sided $t$-test between groups. $P$ values are as follows: i, $2.57 \times 10^{-7}$; ii, $9.21 \times 10^{-8}$.



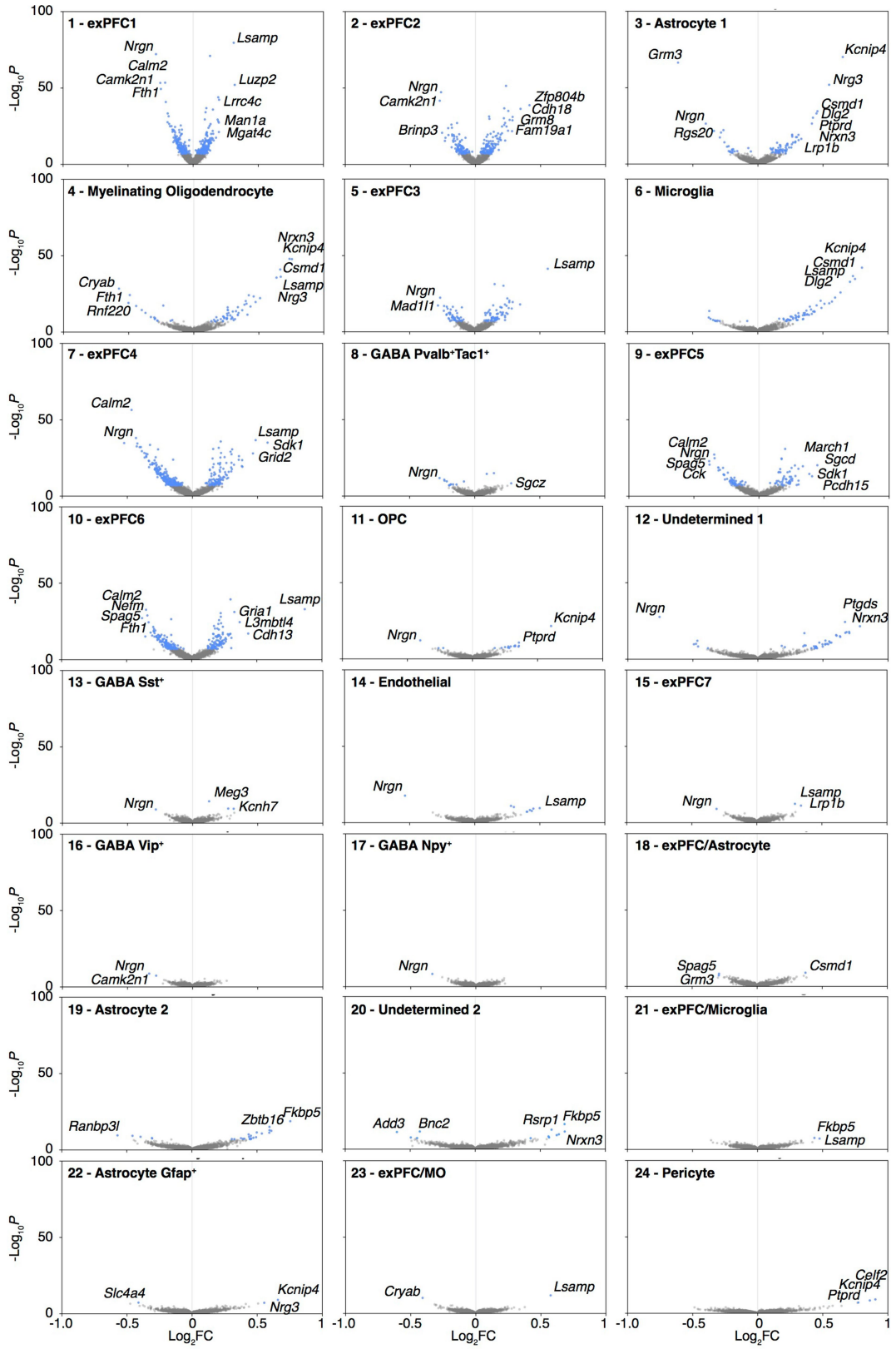

**Extended Data Fig. 3 | Comparable percentages and numbers of CD45$^{high}$ leukocytes in the brains of control and ABX or GF mice. a**, Gating strategy for T cells, B cells, dendritic cells (DCs) and macrophages (Mφ) in the brain. **b**, Population frequencies and numbers of brain-resident CD45$^{high}$ leukocytes in control and ABX mice. **c**, **d**, Population frequencies of CD4$^+$ T cells, CD8$^+$ T cells, CD19$^+$ B cells (**c**), CD11c$^+$ DCs and F4/80$^+$ macrophages (**d**) gated on brain-resident CD45$^{high}$ leukocytes in control and ABX mice. **e**, Population frequencies and numbers of brain-resident CD45$^{high}$ leukocytes in control and GF mice. **f**, **g**, Population frequencies of CD4$^+$ T cells, CD8$^+$ T cells, CD19$^+$ B cells (**f**), CD11c$^+$ DCs and F4/80$^+$ macrophages (**g**) gated on brain-resident CD45$^{high}$ leukocytes in control and GF mice. **h**, Gating strategy of total myeloid cells and Ly6C$^{high}$ monocytes in the brain. **i**, **j**, Population frequencies of total myeloid cells and Ly6C$^{high}$ monocytes gated on brain-resident CD45$^{high}$ leukocytes in control and ABX (**i**) or GF (**j**) mice. Data in **b**, **c**, **g**, **j** are representative of three independent experiments. $n = 4$ mice per group. Data in **d**, **i** are pooled from two independent experiments. $n = 8$ mice per group. Data in **e**, **f** are pooled from two independent experiments. $n = 6$ mice group. Data are mean ± s.e.m.; unpaired two-sided *t*-tests were used. *P* values are indicated on the figures. **k**, Fear extinction in control, GF and *Rag1*$^{-/-}$ mice in the single-session 30-tone fear extinction assay. Data are pooled from two independent experiments. Control $n = 18$, GF $n = 16$, *Rag1*$^{-/-}$ $n = 18$. Mean ± s.e.m.; AUC was calculated for each mouse within each group followed by one-way ANOVA with Tukey's multiple comparisons test. $F(2,49) = 8.558$, $P = 0.0006$. Adjusted *P* values are as follows: i = 0.0343, ii = 0.0004. **l**, Fear extinction of SPF-*Rag1*$^{-/-}$ and GF-*Rag1*$^{-/-}$ mice in the single-session 30-tone fear extinction assay. $n = 7$ mice per group. Mean ± s.e.m.; AUC was calculated for each mouse within each group followed by unpaired two-sided *t*-test between groups. *P* value is shown.



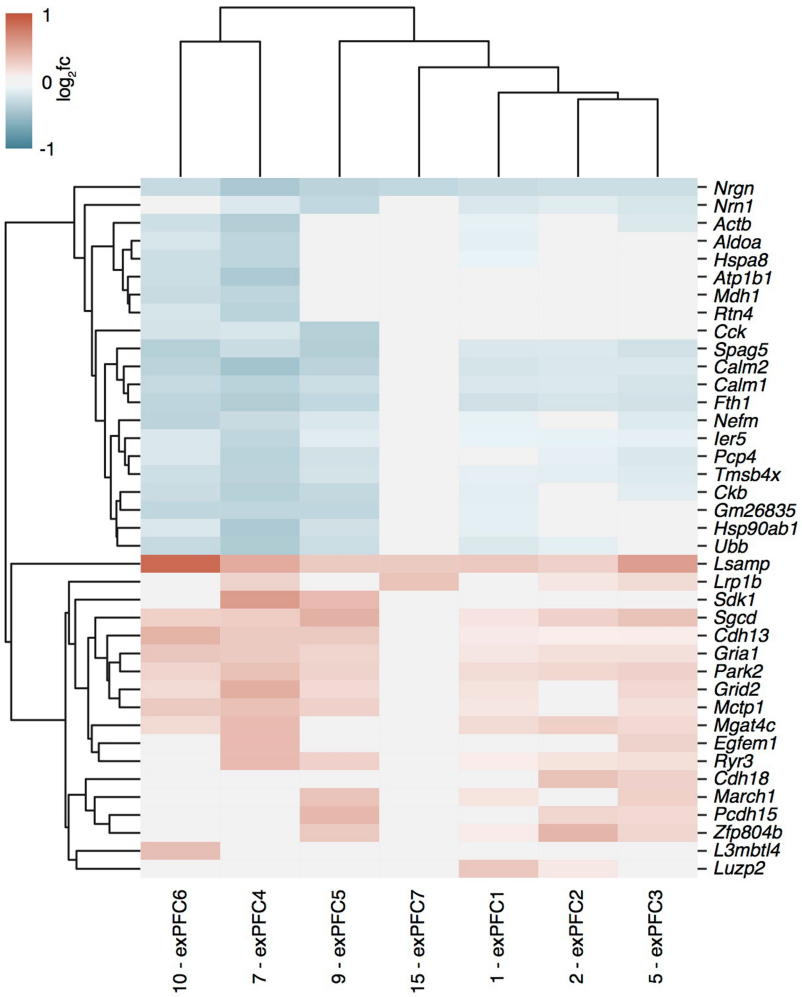

**Extended Data Fig. 4 | Comparable transcriptomes of mPFCs dissected from control and ABX mice in the absence of fear conditioning and extinction. a**, PCA of genome-wide transcriptional profiles of mouse mPFC in the absence of fear conditioning and extinction. Control $n = 3$, ABX $n = 4$. PERMANOVA test was used: $F = 2.52$, Df = 1, $P = 0.17$. **b**, Volcano plot of differential expression between control (negative $\log_2$FC) and ABX (positive $\log_2$FC) groups. DEGs (defined as FDR < 0.1, DESeq2 Wald test) are shown in red. **c**–**f**, Immunofluorescence staining of c-FOS (red) (**c**, **e**) and the density of c-FOS$^+$ neurons (**d**, **f**) in the BLA (**c**, **d**) or IL (**e**, **f**) of control and GF mice 90 min after classical fear extinction session 3. Data pooled from two independent experiments. $n = 6$ mice per group. Mean ± s.e.m.; unpaired two-sided $t$-tests. $P$ values are shown. Scale bar, 200 µm.



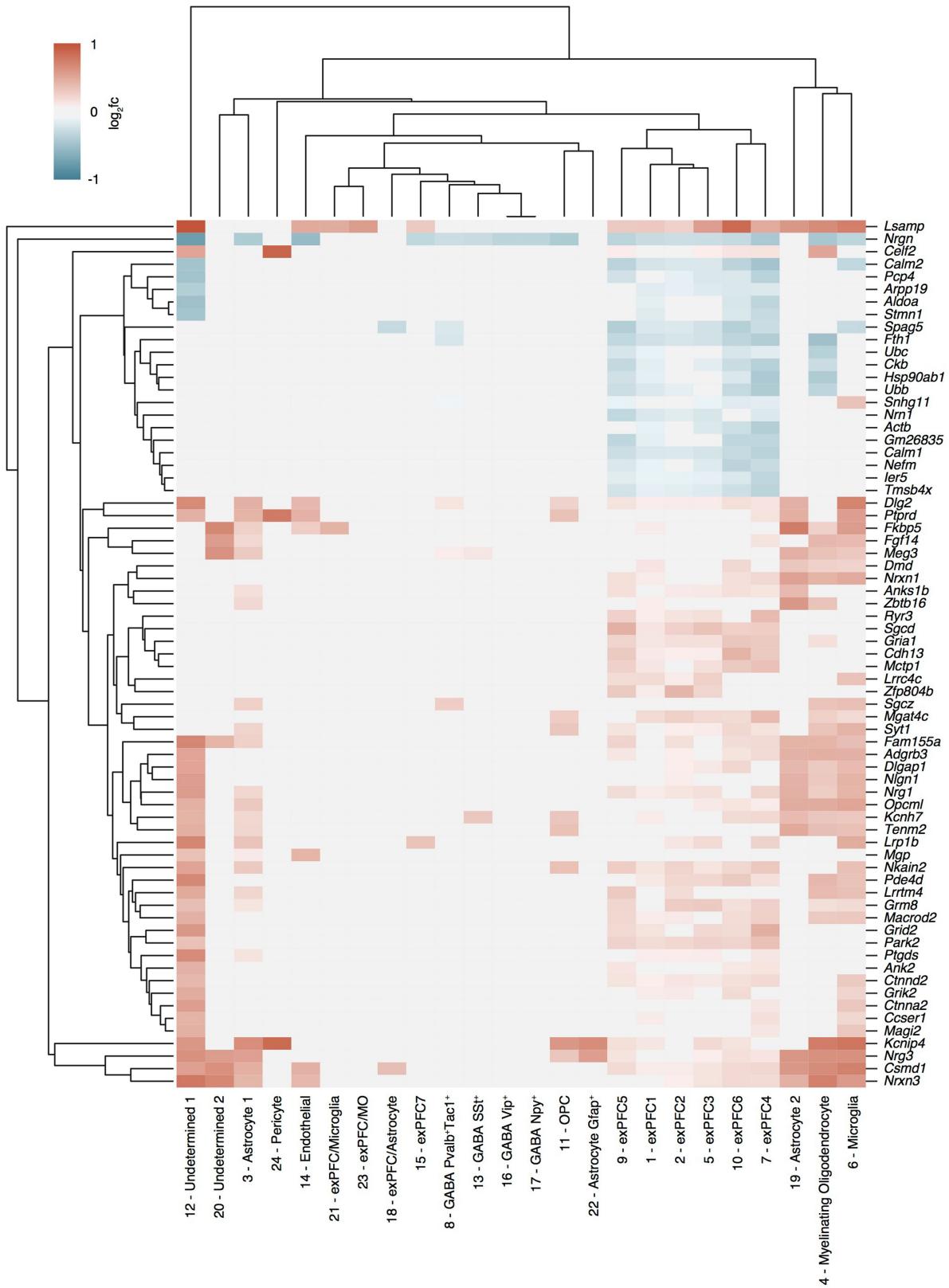

**Extended Data Fig. 5 | Gene expression patterns of individual cell subsets in the mPFC. a**, Proportion of expressing cells (dot size) and mean normalized expression of representative marker genes (columns) associated with the cell clusters shown in Fig. 2a (rows). Clusters are labelled with post facto annotation based on known marker genes. Ambiguous clusters expressing multiple canonical markers across cell types are annotated with both (for example, exPFC/astrocyte), and are likely to represent doublets. **b**, Number of significantly differentially expressed genes (*z*-test calculated on coefficients of mixed linear model, Bonferroni-corrected $P<10^{-7}$) by cluster after downsampling each cluster to 500 nuclei, ranked from highest to lowest (clusters of doublets and undetermined annotations not included). exPFC, glutamatergic excitatory neurons from the PFC; GABA, γ-aminobutyric acid (GABA)ergic interneurons; OPC, oligodendrocyte progenitor cells; MO, myelinating oligodendrocytes.



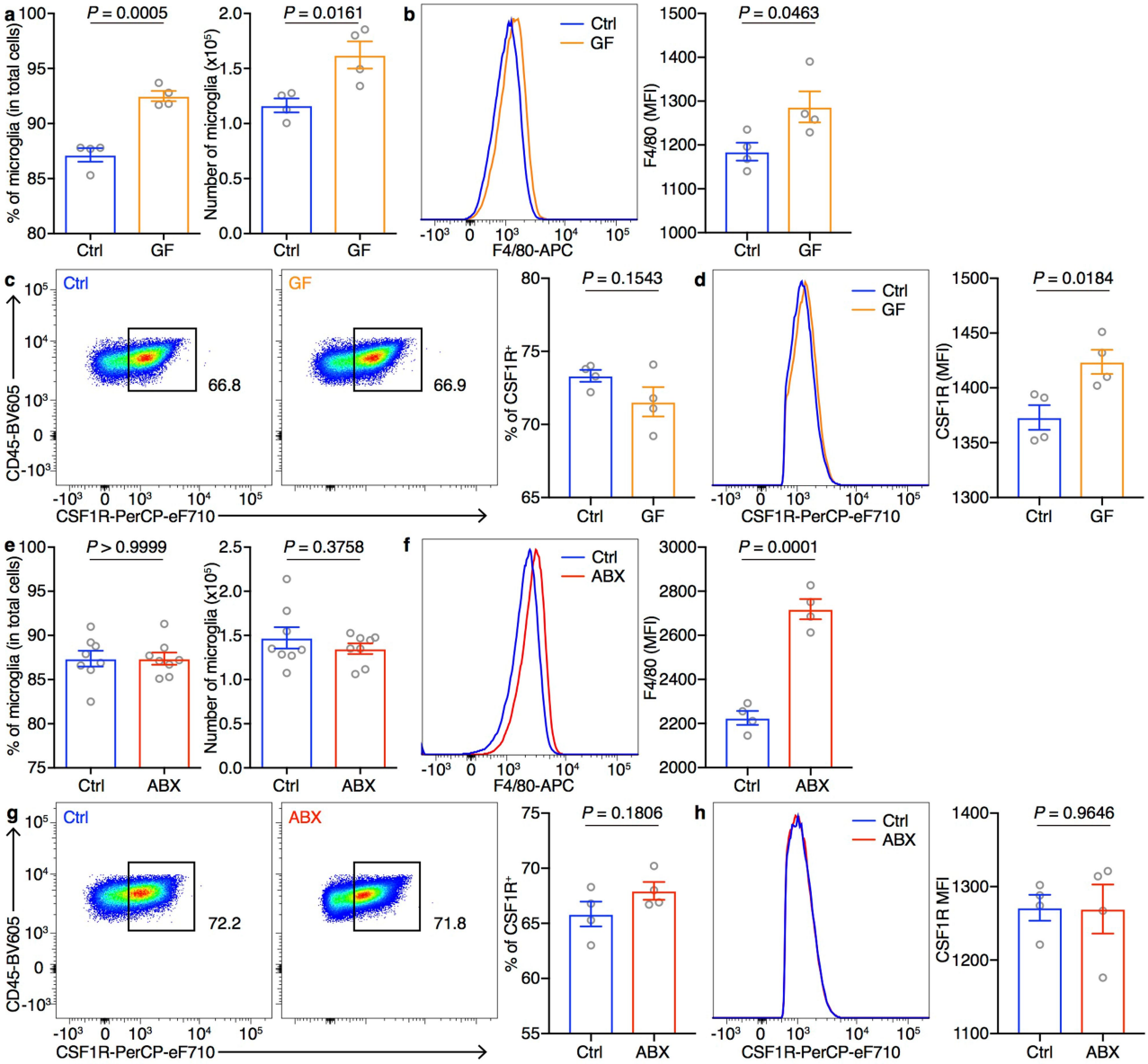

**Extended Data Fig. 6 | Differential gene expression between control and ABX mice in individual clusters of mPFC cells.** Differential expression of ABX versus control ($log_2FC$) in each cluster in Fig. 2a and the associated significance. Blue, genes that are significantly differentially expressed ($z$-test calculated on coefficients of mixed linear model, Bonferroni-corrected $P < 10^{-7}$).



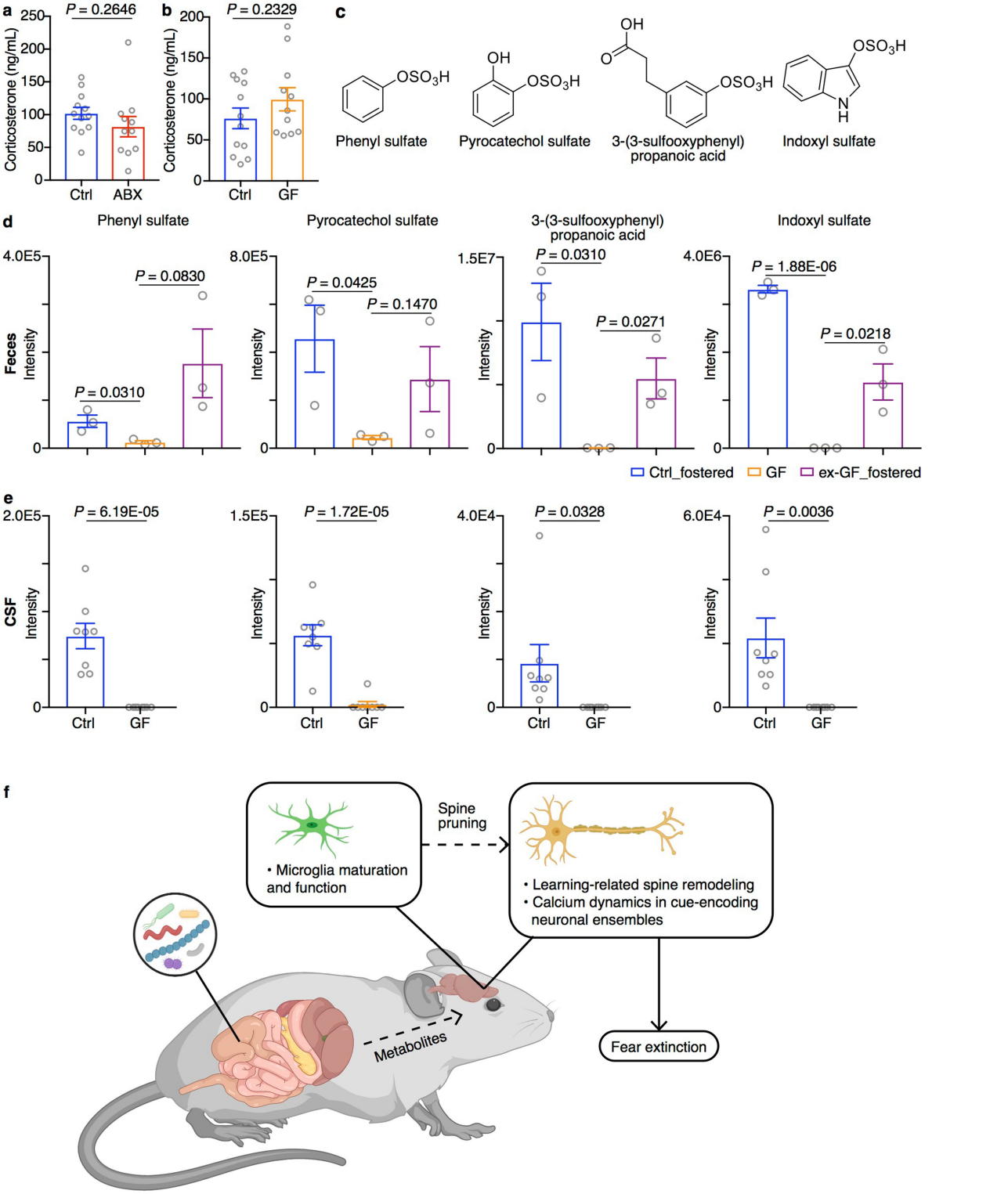

**Extended Data Fig. 7 | Differentially expressed genes in ABX versus control mPFC samples shared by all excitatory neuronal subsets.** Mean fold change in expression in excitatory neurons (columns) from Fig. 2a of genes (rows) that were significantly differentially expressed ($z$-test calculated on coefficients of mixed linear model, Bonferroni-corrected $P < 10^{-7}$) in at least 2 of these clusters, and with absolute $\log_2 FC \geq 0.31$ in at least 1 cluster.




**Extended Data Fig. 8 | Differentially expressed genes in ABX versus control mPFC samples shared by multiple cell types.** Mean fold change in expression across all cell clusters (columns) from Fig. 2a of genes (rows) that were significantly differentially expressed (*z*-test calculated on coefficients of mixed linear model, Bonferroni-corrected $P < 10^{-7}$) in at least 4 clusters, and with absolute $\log_2 FC \geq 0.31$ in at least 1 cluster.




**Extended Data Fig. 9 | Microglia in GF and ABX mice exhibit a developmentally immature phenotype. a**, Population frequencies and numbers of microglia in control and GF mice. **b**, Representative flow cytometry histogram and mean fluorescence intensity (MFI) of F4/80 staining on microglia from control and GF mice. **c**, Representative flow cytometry plots and population frequencies of CSF1R$^+$ microglia in control and GF mice. **d**, Representative flow cytometry histogram and MFI of CSF1R expression gated on CSF1R$^+$ microglia from control and GF mice. Data in **a**–**d** are representative of three independent experiments, $n = 4$ mice per group. **e**, Population frequencies and numbers of microglia in control and ABX mice. **f**, Representative flow cytometry histogram and MFI of F4/80 staining on microglia from control and ABX mice. **g**, Representative flow cytometry plots and population frequencies of CSF1R$^+$ microglia in control and ABX mice. **h**, Representative flow cytometry histogram and MFI of CSF1R expression gated on CSF1R$^+$ microglia from control and ABX mice. Data in **e** are pooled from two independent experiments, $n = 8$ mice per group. Data in **f**–**h** are representative of two independent experiments, $n = 4$ mice per group. Mean ± s.e.m.; unpaired two-sided $t$-test. $P$ values shown.




**Extended Data Fig. 10 | Downregulation of metabolites in GF mice.** **a**, Enzyme-linked immunosorbent assay (ELISA) quantification of plasma corticosterone in control and ABX mice. Data pooled from three independent experiments. Control $n = 12$; ABX $n = 11$. **b**, ELISA quantification of plasma corticosterone in control and GF mice. Data pooled from three independent experiments. Control $n = 12$; GF $n = 11$. Mean ± s.e.m. **c**, Structures of phenyl sulfate, pyrocatechol sulfate, 3-(3-sulfooxyphenyl)propanoic acid and indoxyl sulfate. **d**, Relative abundances of phenyl sulfate, pyrocatechol sulfate, 3-(3-sulfooxyphenyl)propanoic acid and indoxyl sulfate in faecal samples from Ctrl_fostered, GF and ex-GF_fostered mice as determined by LC–MS. $n = 3$ mice per group. **e**, Relative abundances of phenyl sulfate, pyrocatechol sulfate, 3-(3-sulfooxyphenyl)propanoic acid and indoxyl sulfate in CSF samples from control and GF mice as determined by LC–MS. Data are representative of two independent experiments, $n = 8$ mice per group. Mean ± s.e.m.; unpaired two-sided *t*-test. *P* values shown. **f**, Schematic of the microbiota–gut–brain axis in fear extinction learning. Our data inform a model in which alterations in the microbiota and their metabolites influence neuronal function and learning-related plasticity, which may be due to altered microglia-mediated synaptic pruning, and subsequently regulate fear extinction behaviour.

nature research

Corresponding author(s): David Artis, Conor Liston

Last updated by author(s): Aug 7, 2019

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☐ | ☒ | A description of all covariates tested |
| ☐ | ☒ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ | For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☒ | ☐ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☐ | ☒ | Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |

*Our web collection on statistics for biologists contains articles on many of the points above.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | Graphic State v4.0, Stereo Investigator v9, FV31S, ZEN Black v2.6, BD FACSDiva v8.0.1, QuantStudio Real-Time PCR software v1.0 |
| Data analysis | MATLAB R2015a, Prism 7, Fiji, Flowjo 10.4.0, Autotyping 15.04, Image Stabilizer Plugin for ImageJ, CASAVA v1.8.2, FLEXBAR v2.4, STAR v2.3.0, Rsubread, DESeq2 v1.18, vegan R package (https://CRAN.R-project.org/package=vegan), USEARCH v11, phyloseq R package, Cell Ranger v2.0, Scanpy v1.4.0, Louvain algorithm, Python, MSConvert, centWave XCMS algorithm |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

RNA-seq data, 16S rRNA-seq data and single nucleus RNA-seq data are available at Gene Expression Omnibus and BioProject under accession number GSE134808, PRJNA556230 and GSE135326, respectively.

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Sample size | No statistical methods were used to predetermine sample size for experimentation. The sample size was estimated from preliminary experiments or from reports in the literature. |
| Data exclusions | Samples with significant drift during microscopy (i.e. tracked cells or regions that went out of focus or out of frame) were excluded from subsequent analysis. These exclusion criteria were not pre-established though are standard in live timelapse imaging studies. |
| Replication | Experiments were repeated with at least two to three biologically independent for all results presented in the manuscript. If the group size was small (due to limited availability of reagents or mouse strains), data from replicate experiments were pooled for graphical representation. All replicates are biological replicates obtained from biologically independent experiments. |
| Randomization | We did not use randomization to assign animals to experimental groups. As whenever possible littermate controls were used, age did not constitute a variable (and was matched for non-littermates). |
| Blinding | For dendritic spine imaging data analysis, raters blinded to experimental conditions. All other animal studies were not blinded since treatment and experimental analysis could not be separated, blinding of the investigators was not possible. |

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials & experimental systems

| n/a | Involved in the study |
|---|---|
| ☐ | ☒ Antibodies |
| ☒ | ☐ Eukaryotic cell lines |
| ☒ | ☐ Palaeontology |
| ☐ | ☒ Animals and other organisms |
| ☒ | ☐ Human research participants |
| ☒ | ☐ Clinical data |

## Methods

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ ChIP-seq |
| ☐ | ☒ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |

## Antibodies

Antibodies used

The antibodies are described below. All antibodies were purchased from BD, eBioscience (Thermo Fisher), Biolegend, xxx. All antibodies were validated by manufacturers and in previous publications.

Antibodies for flow cytometry:
CD16/CD32 - clone 93 - purified - Biolegend - https://www.biolegend.com/en-us/products/purified-anti-mouse-cd16-32-antibody-190
CD45 - clone 30-F11 -BV605 - Biolegend - https://www.biolegend.com/en-us/products/brilliant-violet-605-anti-mouse-cd45-antibody-8721
CD4 - clone RM4-5 - FITC - eBioscience - https://www.thermofisher.com/antibody/product/CD4-Antibody-clone-RM4-5-Monoclonal/11-0042-82
CD8a - clone 53-6.7 - PE - biolegend - https://www.biolegend.com/en-us/products/pe-anti-mouse-cd8a-antibody-155
CD19 - clone 1D3 - PerCP-Cy5.5 - eBioscience - https://www.thermofisher.com/antibody/product/CD19-Antibody-clone-eBio1D3-1D3-Monoclonal/45-0193-82
CD11b - clone M1/70 - APC-eF780 - eBioscience - https://www.thermofisher.com/antibody/product/CD11b-Antibody-clone-M1-70-Monoclonal/47-0112-82
CD11c - clone N418 - APC-eF780 - eBioscience - https://www.thermofisher.com/antibody/product/CD11c-Antibody-clone-N418-Monoclonal/47-0114-82
F4/80 - clone BM8 - APC - eBioscience - https://www.thermofisher.com/antibody/product/F4-80-Antibody-clone-BM8-Monoclonal/17-4801-82
Ly6G - clone 1A8-Ly6g - PE - eBioscience - https://www.thermofisher.com/antibody/product/Ly-6G-Antibody-clone-1A8-Ly6g-

Monoclonal/12-9668-82
Ly6C - clone HK1.4 - PE-Cy7 - eBioscience - https://www.thermofisher.com/antibody/product/Ly-6C-Antibody-clone-HK1-4-Monoclonal/25-5932-82
CSF1R - clone AFS98 - PerCP-eF710 - eBioscience - https://www.thermofisher.com/antibody/product/CD115-c-fms-Antibody-clone-AFS98-Monoclonal/46-1152-82
All flow antibodies were used at 1:200.

Antibodies for immunofluorescence staining:
c-Fos - clone 4 - Santa Cruz - https://www.scbt.com/scbt/product/c-fos-antibody-4
Synaptophysin - clone SVP-38 - Sigma - https://www.sigmaaldrich.com/catalog/product/sigma/sab4200544?lang=en&region=US
PSD-95 - clone 7E3-1B8 - Sigma - https://www.sigmaaldrich.com/catalog/product/mm/cp35?lang=en&region=US
The above three antibodies were used at 1:1,000.
Donkey anti-Rabbit IgG (H+L) - AF555 - https://www.thermofisher.com/antibody/product/Donkey-anti-Rabbit-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-31572
Donkey anti-Mouse IgG (H+L) - AF555 - https://www.thermofisher.com/antibody/product/Donkey-anti-Mouse-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-31570
The above two secondary antibodies were used at 1:500.

All antibodies information (including catalog number) could be easily found via the vendor websites.

Validation

All antibodies are commercially available and validated by the manufacturer. Vendor websites for antibodies were listed above and the validations can be found there.

## Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

| | |
|---|---|
| Laboratory animals | C57BL/6J (Jax 664), Rag1-/- (Jax 2216), Thy1-YFP-H (Jax 3782) and BALB/c (Jax 651) mice were purchased from The Jackson Laboratory and bred in-house. Male mice were used at 7-16 weeks of age. In individual experiments, all animals were age-matched. All mice were maintained under specific pathogen-free (SPF) conditions on a 12-hour light/dark cycle, and provided food and water ad libitum. Germ-free C57BL/6 mice and gnotobiotic mice were maintained at Weill Cornell Medical College, New York. |
| Wild animals | No wild animals included. |
| Field-collected samples | No field-collected samples included. |
| Ethics oversight | All protocols were approved by the Weill Cornell Medicine Institutional Animal Care and Use Committees (IACUC), and all mice were used in accordance of governmental and institutional guidelines for animal welfare. |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Flow Cytometry

### Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

### Methodology

| | |
|---|---|
| Sample preparation | Sample preparation is described in methods in the 'Brain-resident immune cell isolation and flow cytometry' section. |
| Instrument | A custom configuration Fortessa flow cytometer (BD Biosciences). |
| Software | FACS DIVA software (BD Biosciences) and FlowJo V10 (Tree Star). |
| Cell population abundance | No cell sorting is performed. |
| Gating strategy | The gating strategy is presented in Extended Data Fig. 3 and Extended Data Fig. 9. |

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Article

# Bacterial biodiversity drives the evolution of CRISPR-based phage resistance

https://doi.org/10.1038/s41586-019-1662-9

Received: 21 March 2019

Accepted: 20 September 2019

Published online: 23 October 2019

Ellinor O. Alseth[1]*, Elizabeth Pursey[1], Adela M. Luján[2], Isobel McLeod[1], Clare Rollie[1] & Edze R. Westra[1]*

About half of all bacteria carry genes for CRISPR–Cas adaptive immune systems[1], which provide immunological memory by inserting short DNA sequences from phage and other parasitic DNA elements into CRISPR loci on the host genome[2]. Whereas CRISPR loci evolve rapidly in natural environments[3,4], bacterial species typically evolve phage resistance by the mutation or loss of phage receptors under laboratory conditions[5,6]. Here we report how this discrepancy may in part be explained by differences in the biotic complexity of in vitro and natural environments[7,8]. Specifically, by using the opportunistic pathogen *Pseudomonas aeruginosa* and its phage DMS3vir, we show that coexistence with other human pathogens amplifies the fitness trade-offs associated with the mutation of phage receptors, and therefore tips the balance in favour of the evolution of CRISPR-based resistance. We also demonstrate that this has important knock-on effects for the virulence of *P. aeruginosa*, which became attenuated only if the bacteria evolved surface-based resistance. Our data reveal that the biotic complexity of microbial communities in natural environments is an important driver of the evolution of CRISPR–Cas adaptive immunity, with key implications for bacterial fitness and virulence.

*P. aeruginosa* is a widespread opportunistic pathogen that thrives in a range of different environments, including hospitals, where it is a common source of nosocomial infections. In particular, it frequently colonizes the lungs of patients with cystic fibrosis, in whom it is the leading cause of morbidity and mortality[9]. In part fuelled by a renewed interest in the therapeutic use of bacteriophages as antimicrobials (phage therapy)[10,11], many studies have examined whether and how *P. aeruginosa* evolves resistance to phage (reviewed in ref. [12]). The clinical isolate *P. aeruginosa* strain PA14 has been reported to predominantly evolve resistance against its phage DMS3vir by the modification or complete loss of the phage receptor (type IV pilus) when grown in nutrient-rich medium[5], despite carrying an active CRISPR–Cas adaptive immune system. By contrast, under nutrient-limited conditions, the same strain relies on CRISPR–Cas to acquire phage resistance[5]. These differences are due to higher phage densities during infections in nutrient-rich compared with nutrient-limited conditions, which in turn determines whether surface-based resistance (with a fixed cost of resistance) or CRISPR-based resistance (infection-induced cost) is favoured by natural selection[5,13]. Although these observations suggest abiotic factors are crucial determinants of the evolution of phage resistance strategies, the role of biotic factors has remained unclear, even though *P. aeruginosa* commonly co-exists with a range of other bacterial species in both natural and clinical settings[14,15]. We proposed that the presence of a bacterial community could drive increased levels of CRISPR-based resistance evolution for two main reasons. First, reduced densities of *P. aeruginosa* in the presence of competitors may limit phage amplification, and favour CRISPR-based resistance[5]. Second, pleiotropic costs associated with the mutation of phage receptors may be amplified during interspecific competition.

## Bacterial biodiversity drives CRISPR evolution

To explore these hypotheses, we co-cultured *P. aeruginosa* PA14 with three other clinically relevant opportunistic pathogens that can co-infect with *P. aeruginosa*, namely *Staphylococcus aureus*, *Burkholderia cenocepacia* and *Acinetobacter baumannii*[14–17], none of which can be infected by or interact with phage DMS3vir (Extended Data Fig. 1). We applied a 'mark–recapture' approach using a *P. aeruginosa* PA14 mutant carrying streptomycin resistance to monitor the bacterial population dynamics and evolution of phage resistance in the focal subpopulation at 3 days post infection (d.p.i.). This revealed that in nutrient-rich lysogeny broth, PA14 evolved significantly higher levels of CRISPR-based resistance after infection with $10^6$ plaque-forming units (p.f.u.) of phage DMS3vir when co-cultured with other bacterial species than when grown in isolation or co-cultured with an isogenic surface mutant (Fig. 1a). In addition, we found that these effects were dependent on the identity of the species that were present in the mixed culture, with the strongest effects being observed in the presence of *A. baumannii* or a mix of the three bacterial species, and an absence of any effect when PA14 was co-cultured with an isogenic surface mutant that lacked the phage receptor (Fig. 1a, deviance test: relationship between community composition and CRISPR; residual deviance (30, $n = 36$) = 1.81, $P = 2.2 \times 10^{-16}$; Tukey contrasts: monoculture versus mixed; $z = -5.99$, $P = 3.02 \times 10^{-8}$; monoculture versus *A. baumannii*; $z = -4.33$, $P = 0.00023$; monoculture versus

[1]Environment and Sustainability Institute, Biosciences, University of Exeter, Penryn Campus, Penryn, UK. [2]IRNASUS, CONICET, Facultad de Ciencias Químicas, Universidad Católica de Córdoba, Córdoba, Argentina. *e-mail: eao210@exeter.ac.uk; e.r.westra@exeter.ac.uk

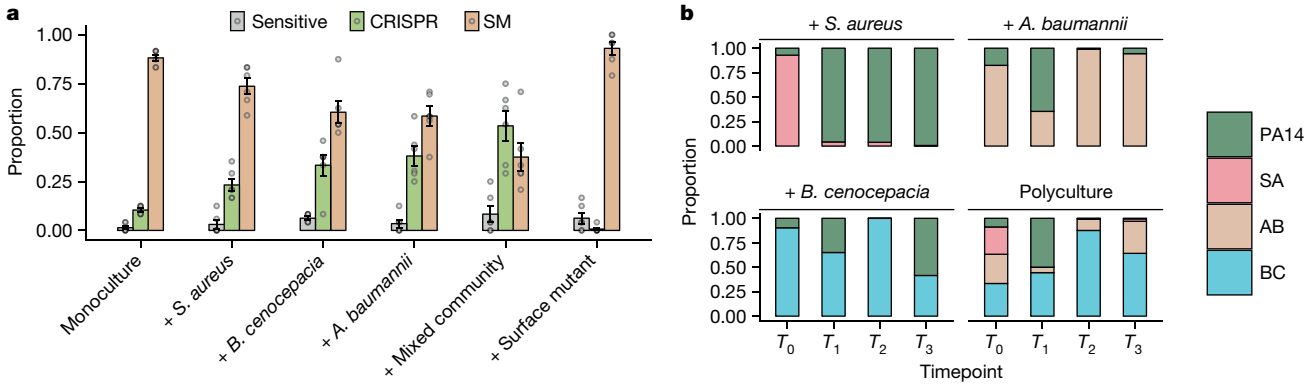



**Fig. 1 | Biodiversity affects the evolution of phage resistance. a**, Proportion of *P. aeruginosa* that acquired surface modification (SM) or CRISPR-based resistance, or remained sensitive at 3 d.p.i. with phage DMS3vir when grown in monoculture or polycultures, or with an isogenic surface mutant (6 replicates per treatment, with 24 colonies per replicate, $n = 36$ biologically independent replicates). Data are mean ± s.e.m. **b**, Microbial community composition over time for the mixed-species infection experiments. AB, *A. baumannii*; BC, *B. cenocepacia*; PA14, *P. aeruginosa*; SA, *S. aureus*.

*B. cenocepacia*; $z = -3.76$, $P = 0.0026$; monoculture versus *S. aureus*; $z = -2.38$, $P = 0.26$; monoculture versus surface mutant; $z = 2.26$, $P = 0.35$). Notably, *P. aeruginosa* densities were strongly reduced in the presence of *A. baumannii*, *B. cenocepacia* and the mixed community, whereas *P. aeruginosa* dominated the community during competition with *S. aureus* despite the presence of phage DMS3vir (Fig. 1b), which suggests a positive relationship between the strength of interspecific competition and the levels of CRISPR-based resistance evolution.

Next, to explore the potential clinical relevance of this observation, we performed a similar experiment in artificial sputum medium (ASM), which is a nutrient-rich medium that mimics the abiotic environment of sputum from patients with cystic fibrosis[18]. This revealed a similar pattern as that observed in lysogeny broth, with *A. baumannii* and the community as a whole resulting in a marked increase in CRISPR-based resistance (Extended Data Fig. 2). To explore the generality of these findings further, we also manipulated the composition of the microbial community by varying the proportion of *P. aeruginosa* versus the other pathogens. This revealed that increased CRISPR-based resistance evolution occurred across a wide range of microbial community compositions, with a maximum effect size when *P. aeruginosa* made up 50% of the initial mixture (Extended Data Fig. 3). An exception to this trend was when the *P. aeruginosa* subpopulation made up only 1% of the total community; in this case, sensitive bacteria persisted alongside resistant bacteria because of the reduced size of the phage epidemic and hence relaxed selection for resistance (Extended Data Fig. 3). Collectively, these data suggest that greater levels of interspecific competition contribute to the evolution of CRISPR-based resistance.

## Biodiversity amplifies costs of surface resistance

We hypothesized that reduced population sizes of *P. aeruginosa* in the presence of competitors might explain the increased evolution of CRISPR-based resistance, as this leads to smaller phage epidemics, which is known to favour CRISPR-based over surface-based resistance[5]. However, variation in the force of infection did not seem to have a strong role in the observed effects, because even though phage epidemic sizes varied depending on the microbial community composition (Extended Data Fig. 4), this did not correlate with the levels of evolved CRISPR resistance (Extended Data Fig. 5). Moreover, when manipulating the starting titres of the DMS3vir phage, we observed no differences in the levels of evolved CRISPR-based resistance when *P. aeruginosa* was co-cultured in the presence of the microbial community (Extended Data Fig. 6). An alternative explanation for the observed effects may therefore be that the fitness cost of surface-based resistance is amplified in the presence of other bacterial species; for example, owing to cell-surface molecules playing a part in interspecific competition[19], which again would result in stronger selection towards bacteria with CRISPR-based resistance. To test this hypothesis, we competed the two phage-resistant phenotypes (that is, CRISPR-resistant and surface mutant) in the presence or absence of the microbial community, and across a range of phage titres. In the absence of the microbial community and phage, CRISPR-resistant bacteria had a small fitness advantage over bacteria with surface-based resistance, but this advantage disappeared when phage was added and as titres increased[5] (Fig. 2a). In the presence of the biodiverse microbial community, however, the relative fitness of bacteria with CRISPR-based resistance was consistently higher, which demonstrates that mutation of the type IV pilus is more costly when bacteria compete with other bacterial species (Fig. 2a, linear model: effect of community absence; $t = -5.54$, $P = 1.49 \times 10^{-7}$; effect of increasing phage titre; $t = -2.41$, $P = 0.017$; overall model fit; adjusted $R^2 = 0.41$, $F_{4,139} = 25.48$, $P = 7.65 \times 10^{-16}$). The increased fitness trade-off associated with surface-based resistance was also observed when the CRISPR- and surface-resistant phenotypes competed in the presence of only a single additional species (Fig. 2b, two-way ANOVA with Tukey contrasts: overall difference in fitness; $F_{4,2} = 8.151$ $P = 6.31 \times 10^{-6}$; monoculture versus mixed; $P = 0.011$; monoculture versus *A. baumannii*; $P = 0.016$; monoculture versus *B. cenocepacia*; $P = 0.022$), with the exception of *S. aureus* (Fig. 2d, monoculture versus *S. aureus*; $P = 0.80$), concordant with this species being the weakest competitor and inducing the lowest levels of CRISPR-based resistance (Fig. 1). These fitness trade-offs therefore explain why *P. aeruginosa* evolved greater levels of CRISPR-based resistance in the presence of the other pathogens, and why this varied depending on the competing species (Fig. 1).

## CRISPR-resistant *P. aeruginosa* remains virulent

The evolution of phage resistance by bacterial pathogens is often associated with virulence trade-offs when surface structures are modified[20], whereas similar trade-offs have not yet been reported for CRISPR-based resistance. We therefore hypothesized that the community context in which phage resistance evolves may have important knock-on effects for *P. aeruginosa* virulence. To test this, we used a *Galleria mellonella* infection model, which is commonly used to evaluate the virulence of human pathogens[21,22]. We compared the in vivo virulence of *P. aeruginosa* clones that evolved phage resistance against phage DMS3vir in different community contexts by injecting larvae with a mixture of clones that had evolved phage resistance in either the presence or absence of the mixed bacterial community (Extended Data Fig. 3c). By taking the time until death as a proxy for virulence, we found that the evolution of phage resistance in the presence of a microbial community was associated with greater levels of *P. aeruginosa* virulence than when phage-resistance

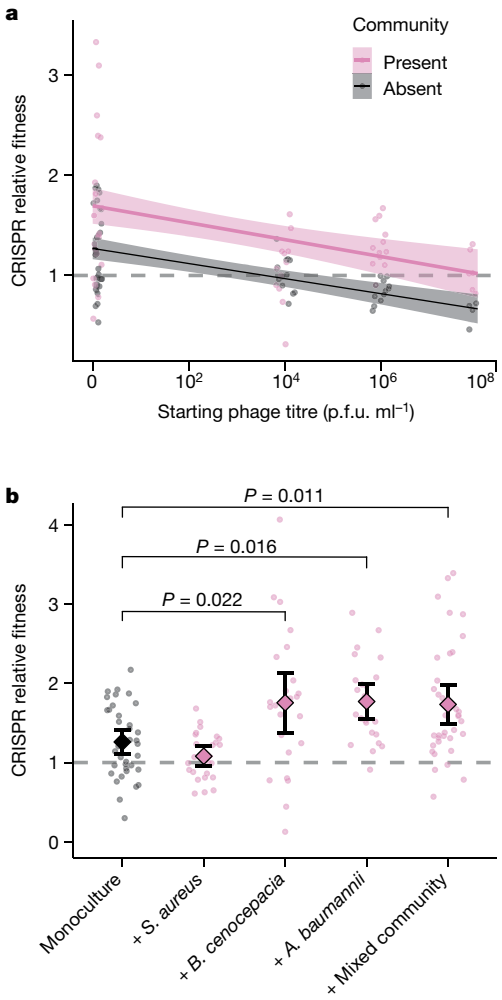



**Fig. 2 | Biodiversity amplifies fitness costs associated with surface-based resistance. a**, Relative fitness of a *P. aeruginosa* clone with CRISPR-based resistance after competing for 24 h against a surface-modification clone at varying titres of phage DMS3vir in the presence or absence of a mixed microbial community. Regression slopes with shaded areas corresponding to 95% confidence interval ($n = 144$ biologically independent samples). **b**, Relative fitness after competition in the absence of phage, but in the presence of other bacterial species individually or as a mixture. Data are mean and 95% confidence intervals ($n = 144$ biologically independent samples).

evolved in monoculture, and remained similar to that of the ancestral PA14 strain (Fig. 3a, Cox proportional-hazards model with Tukey contrasts: community present versus absent; $z = 5.85$, $P = 1 \times 10^{-4}$; ancestral PA14 versus community absent; $z = 4.42$, $P = 1 \times 10^{-4}$; ancestral PA14 versus community present; $z = -1.30$, $P = 0.38$; overall model fit; $LRT_3 = 51.03$, $n = 376$, $P = 5 \times 10^{-11}$). These data, in combination with the fact that the type IV pilus is a well-known virulence factor[23], are consistent with the notion that the mechanism by which bacteria evolve phage resistance has important implications for bacterial virulence. To test this more directly, we next infected larvae with each individual *P. aeruginosa* clone for which we had previously determined the mechanism that underlies evolved phage resistance (Extended Data Fig. 3c), again using the time until death as a measure of virulence. This showed that bacterial clones with surface-based resistance—unlike those with CRISPR-based resistance—both had markedly reduced swarming motility (as expected for mutations in the type IV pilus[23]) (Fig. 3b; one-way ANOVA with Tukey contrasts: overall effect; $F_{2,977} = 472.5$, $P = 2.2 \times 10^{-16}$; sensitive versus CRISPR; $P = 0.87$; CRISPR versus surface mutant; $P = 1 \times 10^{-5}$) and impaired virulence compared with phage-sensitive bacteria (Fig. 3c; Cox proportional-hazards model with Tukey contrasts: surface mutant versus CRISPR; $z = -2.37$, $P = 0.045$; sensitive versus CRISPR; $z = 2.10$, $P = 0.10$; surface mutant versus sensitive; $z = -4.23$, $P = 1 \times 10^{-3}$; overall model fit; $LRT_3 = 48.66$, $n = 981$, $P = 2 \times 10^{-10}$). Similar virulence trade-offs were also observed when larvae were injected with *P. aeruginosa* PA14 clones that had evolved surface-based resistance against phage LMA2, which uses lipopolysaccharide (LPS) as a receptor (Extended Data Fig. 7).

## Discussion

We have shown that the evolutionary outcome of bacteria–phage interactions can be fundamentally altered by the microbial community context. Although conventionally studied in isolation, these interactions are usually embedded in complex biotic networks of interacting species, and it is becoming increasingly clear that this can have key implications for the evolutionary epidemiology of infectious disease[24–28]. Our work shows that the community context can shape the evolution of different host-resistance strategies. Specifically, we find that the interspecific interactions between four bacterial species in a synthetic microbial community can have a large effect on the evolution of phage-resistance mechanisms by amplifying the constitutive fitness cost of surface-based resistance[5]. The finding that biotic complexity matters complements previous work on the effect of abiotic variables and force of infection on the evolution of phage resistance[5]. The data presented here suggest that the effect of biotic complexity on the evolution of CRISPR-based resistance is stronger than that of variation in phage abundance, which is consistent with the observation that in the presence of the polymicrobial

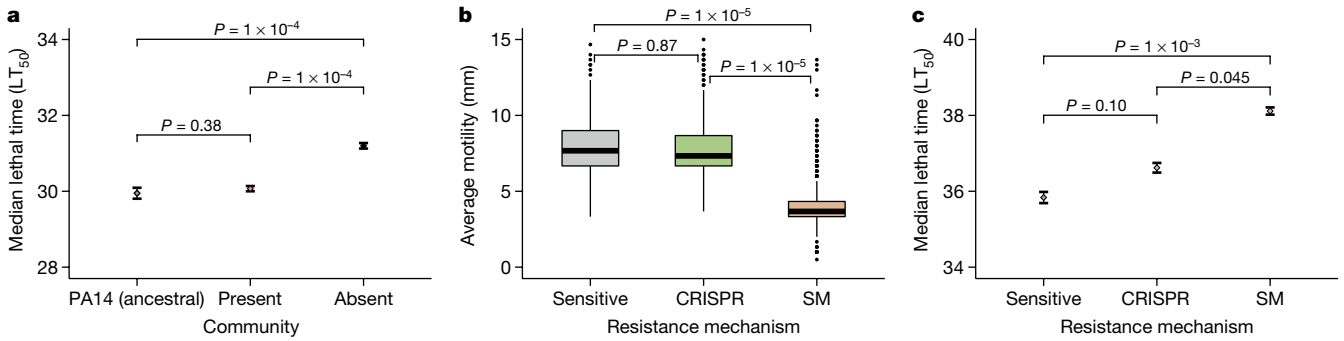



**Fig. 3 | Evolution of phage resistance affects in vivo virulence. a**, Time until death (given as the median ± one standard error) after infection with PA14 clones that evolved phage resistance in either the presence or the absence of a mixed microbial community ($n = 376$ biologically independent samples, analysed using a Cox proportional-hazards model with Tukey contrasts). $LT_{50}$, median lethal time. **b**, The effect of the type of evolved phage resistance (CRISPR-based or surface-modification-based) on bacterial motility ($n = 981$ biologically independent samples). Box plots show the median with the upper and lower twenty-fifth and seventy-fifth percentiles, the interquartile range, and outliers shown as dots. **c**, The effect of the type of resistance on in vivo virulence (time until death, given as the median ± one standard error; $n = 981$, analysed using a Cox proportional-hazards model with Tukey contrasts).

community, bacteria with CRISPR-based resistance outcompeted bacteria with surface-based resistance at all phage titres (Fig. 2). The amplified fitness cost of surface mutation also suggests that the type IV pilus has an important role in interspecific competition. Although future work will be crucial to understand the detailed molecular mechanism that underpins these effects, and to generalize the findings described here to other bacterial species and strains, we speculate that the way in which the composition of the microbial community drives the evolution of phage-resistance strategies may be important in the context of phage therapy. Primarily, the absence of detectable trade-offs between CRISPR-based resistance and virulence, as opposed to when bacteria evolve surface-based resistance, suggests that the evolution of CRISPR-based resistance can ultimately influence the severity of disease. Moreover, the evolution of CRISPR-based resistance can drive more rapid phage extinction[29], and may in a multi-phage environment result in altered patterns of cross-resistance evolution compared with surface-based resistance[30]. The identification of the drivers and consequences of CRISPR-resistance evolution might help to improve our ability to predict and manipulate the outcome of bacteria–phage interactions in both natural and clinical settings.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1662-9.


1. Grissa, I., Vergnaud, G. & Pourcel, C. CRISPRcompar: a website to compare clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* **36**, W145–W148 (2008).
2. Barrangou, R. et al. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–1712 (2007).
3. Andersson, A. F. & Banfield, J. F. Virus population dynamics and acquired virus resistance in natural microbial communities. *Science* **320**, 1047–1050 (2008).
4. Laanto, E., Hoikkala, V., Ravantti, J. & Sundberg, L. R. Long-term genomic coevolution of host–parasite interaction in the natural environment. *Nat. Commun.* **8**, 111 (2017).
5. Westra, E. R. et al. Parasite exposure drives selective evolution of constitutive versus inducible defense. *Curr. Biol.* **25**, 1043–1049 (2015).
6. van Houte, S., Buckling, A. & Westra, E. R. Evolutionary ecology of prokaryotic immune mechanisms. *Microbiol. Mol. Biol. Rev.* **80**, 745–763 (2016).
7. Hibbing, M. E., Fuqua, C., Parsek, M. R. & Peterson, S. B. Bacterial competition: surviving and thriving in the microbial jungle. *Nat. Rev. Microbiol.* **8**, 15–25 (2010).
8. O'Toole, G. A. Cystic fibrosis airway microbiome: overturning the old, opening the way for the new. *J. Bacteriol.* **200**, 1–8 (2018).
9. Folkesson, A. et al. Adaptation of *Pseudomonas aeruginosa* to the cystic fibrosis airway: an evolutionary perspective. *Nat. Rev. Microbiol.* **10**, 841–851 (2012).
10. Roach, D. R. & Debarbieux, L. Phage therapy: awakening a sleeping giant. *Emerg. Top. Life Sci.* **1**, 93–103 (2017).
11. Rossitto, M., Fiscarelli, E. V. & Rosati, P. Challenges and promises for planning future clinical research into bacteriophage therapy against *Pseudomonas aeruginosa* in cystic fibrosis. An argumentative review. *Front. Microbiol.* **9**, 775 (2018).
12. De Smet, J., Hendrix, H., Blasdel, B. G., Danis-Wlodarczyk, K. & Lavigne, R. *Pseudomonas* predators: understanding and exploiting phage–host interactions. *Nat. Rev. Microbiol.* **15**, 517–530 (2017).
13. Chabas, H., van Houte, S., Høyland-Kroghsbo, N. M., Buckling, A. & Westra, E. R. Immigration of susceptible hosts triggers the evolution of alternative parasite defence strategies. *Proc. R. Soc. B* **283**, 20160721 (2016).
14. Harrison, F. Microbial ecology of the cystic fibrosis lung. *Microbiology* **153**, 917–923 (2007).
15. O'Brien, S. & Fothergill, J. L. The role of multispecies social interactions in shaping *Pseudomonas aeruginosa* pathogenicity in the cystic fibrosis lung. *FEMS Microbiol. Lett.* **364**, 1–10 (2017).
16. Bhargava, N., Sharma, P. & Capalash, N. Pyocyanin stimulates quorum sensing-mediated tolerance to oxidative stress and increases persister cell populations in *Acinetobacter baumannii*. *Infect. Immun.* **82**, 3417–3425 (2014).
17. Rocha, G. A. et al. Species distribution, sequence types and antimicrobial resistance of *Acinetobacter* spp. from cystic fibrosis patients. *Epidemiol. Infect.* **146**, 524–530 (2018).
18. Diraviam Dinesh, S. & Diraviam Dinesh, S. Artificial sputum medium. *Protoc. Exchange* https://doi.org/10.1038/protex.2010.212 (2010).
19. An, D., Danhorn, T., Fuqua, C. & Parsek, M. R. Quorum sensing and motility mediate interactions between *Pseudomonas aeruginosa* and *Agrobacterium tumefaciens* in biofilm cocultures. *Proc. Natl Acad. Sci. USA* **103**, 3828–3833 (2006).
20. León, M. & Bastías, R. Virulence reduction in bacteriophage resistant bacteria. *Front. Microbiol.* **6**, 343 (2015).
21. Kavanagh, K. & Reeves, E. P. Exploiting the potential of insects for in vivo pathogenicity testing of microbial pathogens. *FEMS Microbiol. Rev.* **28**, 101–112 (2004).
22. Hernandez, R. J. et al. Using the wax moth larva *Galleria mellonella* infection model to detect emerging bacterial pathogens. *PeerJ* **6**, e6150 (2019).
23. Craig, L., Pique, M. E. & Tainer, J. A. Type IV pilus structure and bacterial pathogenicity. *Nat. Rev. Microbiol.* **2**, 363–378 (2004).
24. Johnson, P. T. J., de Roode, J. C. & Fenton, A. Why infectious disease research needs community ecology. *Science* **349**, 1259504 (2015).
25. Alizon, S., de Roode, J. C. & Michalakis, Y. Multiple infections and the evolution of virulence. *Ecol. Lett.* **16**, 556–567 (2013).
26. Benmayor, R., Hodgson, D. J., Perron, G. G. & Buckling, A. Host mixing and disease emergence. *Curr. Biol.* **19**, 764–767 (2009).
27. Keesing, F. et al. Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature* **468**, 647–652 (2010).
28. Chabas, H. et al. Evolutionary emergence of infectious diseases in heterogeneous host populations. *PLoS Biol.* **16**, e2006738 (2018).
29. van Houte, S. et al. The diversity-generating benefits of a prokaryotic adaptive immune system. *Nature* **532**, 385–388 (2016).
30. Wright, R. C. T., Friman, V. P., Smith, M. C. M. & Brockhurst, M. A. Cross-resistance is modular in bacteria–phage interactions. *PLoS Biol.* **16**, e2006057 (2018).

# Methods

All statistical analyses were done using R v.3.5.1[33], and the Tidyverse package v.1.2.1[34]. All *G. mellonella* mortality analyses were done using the Survival package v.2.38[35]. No statistical methods were used to predetermine sample size. The experiments were not randomized, and investigators were not blinded to allocation during experiments and outcome assessment.

### Bacterial strains and viruses

We used a marked *P. aeruginosa* UCBPP-PA14 mutant carrying a streptomycin-resistant gene inserted into the genome using pBAM1[31] (referred to as the ancestral PA14 strain). The wild-type PA14 bacteriophage-insensitive mutant with two CRISPR spacers (BIM2), the surface mutant derived from the PA14 *csy3::LacZ* strain, and phage DMS3vir and DMS3vir+*acrF1* (carrying an anti-CRISPR gene) have all been previously described (refs. [5,29] and references therein). The bacteria used as the microbial community were *S. aureus* strain 13 S44 S9, *A. baumannii* clinical isolate FZ21 and *B. cenocepacia* J2315, and were all isolated from patients at Queen Astrid Military Hospital, Brussels, Belgium.

### Adsorption and infection assays

Phage infectivity against each of the bacterial species used in this study was assessed by spotting serial dilutions of virus DMS3vir on lawns of the individual community bacteria, followed by checking for any plaque formation after 24 h of growth at 37 °C. Adsorption assays (as shown in Extended Data Fig. 1) were performed by monitoring phage titres over time, for up to an hour (at 0, 2, 4, 6, 8, 10, 15 and 20 min after infection for PA14, and at 0, 5, 10, 20, 40 and 60 min after infection for the other bacteria species; for the no-bacterial control, sampling was done at 0 and 60 min after infection), after inoculating the individual bacteria in mid-log phase at approximately $2 \times 10^8$ c.f.u. with phage DMS3vir at $2 \times 10^6$ p.f.u. (final multiplicity of infection = 0.001). Adsorption assays were carried out in falcon tubes containing 15 ml LB medium, incubated at 37 °C while shaking at 180 r.p.m. (three independent replicates per experiment). At each time point, 50 µl of sample was transferred to pre-cooled Eppendorfs on ice, containing 900 µl LB medium and 50 µl chloroform, before vortexing for 10 s. After sampling was completed, all eppendorfs were centrifuged at 13,000 r.p.m. at 4 °C for at least 5 min after which 300 µl of the supernatant was extracted, diluted and spotted onto lawns of *P. aeruginosa* followed by checking for plaque formation after 24 h of growth at 37 °C.

### Evolution experiments

The streptomycin-resistant mutant of the ancestral strain of *P. aeruginosa* was used for all evolution experiments. Evolution experiments (shown in Fig. 1 and Extended Data Figs. 2, 3) were performed by inoculating 60 µl from overnight cultures (containing approximately $10^6$ c.f.u.) into glass microcosms containing 6 ml LB medium (Fig. 1 and Extended Data Fig. 3), or ASM[18] (Extended Data Fig. 2). One litre of ASM was made by mixing 5 g mucin from porcine stomach (Sigma), 4 g low molecular-mass salmon sperm DNA (Sigma), 5.9 mg diethylene triamine pentaacetic acid (DTPA) (Sigma), 5 g NaCl (Sigma), 2.2 g KCl (Sigma), 1.81 g Tris base (Thermo Fisher Scientific), 5 ml egg yolk emulsion (Sigma) and 250 mg of each of 20 amino acids (Sigma), as previously described[18]. Inoculation was followed by incubation at 37 °C while shaking at 180 r.p.m. ($n = 6$ per treatment). The polyculture mixes either consisted of approximately equal amounts of all four bacterial species or mixes of *P. aeruginosa* with just one additional species where *P. aeruginosa* made up 25% of the total volume used for inoculation (that is, 15 µl of 60 µl), unless otherwise indicated (Extended Data Fig. 3). Before inoculation, phage DMS3vir was added at $10^6$ p.f.u. (Fig. 1 and Extended Data Fig. 2), or at $10^4$ p.f.u. (Extended Data Fig. 3). Transfers of 1:100 into fresh broth were done daily for a total of three days. In addition, phage titres were monitored daily by spotting chlorophorm-treated lysate dilutions on a lawn of *P. aeruginosa csy3::LacZ*. Downstream analysis to determine whether and how bacteria evolved phage resistance was done by cross-streak assays and PCR on 24 randomly selected clones per replicate experiment, as previously described[5].

### DNA extraction and qPCR

For the experiment shown in Fig. 1, the densities of the different bacterial species in the microbial communities over time were determined using qPCR. DNA was extracted from all replicas using the DNeasy UltraClean Microbial Kit (Qiagen), following the manufacturer's instructions. Before DNA extraction, to ensure lysis of *S. aureus*, 15 µl lysostaphin (Sigma) at 0.1 mg ml$^{-1}$ was added to 500 µl of sample followed by incubation at 37 °C for at least 1 h. For *P. aeruginosa*, *A. baumannii* and *B. cenocepacia*, the 16S gene was chosen as the target for the qPCR primers and were as follows: PA14 forward primer (PA14-16 s-F), AGTTGGGAGGAAGGGCAGTA; PA14 reverse primer (PA14-16 s-R), GCTTGCTGAACCACTTACGC; *A. baumannii* forward primer (AB-16 s-F), ATCAGAATGCCGCGGTGAAT; *A. baumannii* reverse primer (AB-16 s-R), ACCGCCCTCTTTGCAGTTAG; *B. cenocepacia* forward primer (BC-16 s-F), ATACAGTCGGGGGATGACGG; *B. cenocepacia* reverse primer (BC-16 s-R), TCACCAATGCAGTTCCCAGG. For *S. aureus*, we used qPCR primers that have previously been described[32]. The amplification reactions were performed in triplicates, with Brilliant SYBR Green reagents (Agilent) in 20 µl reactions made up of 10 µl master mix, 2 µl primer pair, 0.4 µl dye, and sterile nuclease free water to a total volume of 15 µl before adding 5 µl diluted DNA sample. The qPCR program was as follows: 95 °C for 3 min, 40 cycles at 95 °C for 10 s and 60 °C for 30 s. All qPCR reactions and results were analysed using the Applied Biosystems QuantStudio 7 Flex Real-Time PCR system.

### Competition experiments

For both competition experiments in Fig. 2, the BIM2 clone was competed against the surface mutant derived from the PA14 *csy3::LacZ* strain[5]. Bacteria were grown for 24 h in glass microcosms containing 6 ml LB medium, in a shaking incubator at 180 r.p.m. and at 37 °C. For the experiment in Fig. 2a, the two phenotypes were competed in the presence or absence of the mixed microbial community, either without the addition of phage ($n = 36$), or infected with phage DMS3vir at $10^4$, $10^6$ and $10^8$ p.f.u. ($n = 12$ per treatment). For the experiment shown in Fig. 2b, the two phage-resistant phenotypes were again competed either in the presence or absence of individual bacterial species or a mixed community of all species. *P. aeruginosa* made up 25% of the total volume of 60 µl that was used to inoculate the 6 ml of LB medium ($n = 24$ per treatment). Samples were taken at 0 and 24 h after infection, and the cells were serial diluted in M9 salts and plated on cetrimide agar (Sigma) supplemented with approximately 50 µg ml$^{-1}$ X-gal (to select for *P. aeruginosa*, while also differentiating between the CRISPR-resistant clones (white) and the surface mutant (blue)). Relative fitness was calculated as previously described[5,29].

### Virulence assays

All infection experiments were done using *G. mellonella* larvae (UK WaxWorms Ltd). Throughout the experiments, the larvae were stored in 12-well plates, with one larva per well, and were all checked for mortality and melanization before injection. Bacterial inoculums were prepared depending on the experiment, and were as follows. For the experiment in Fig. 3a, all 24 evolved clones from each replicate from the 25% (community present) and 100% (community absent) treatments (Extended Data Fig. 3) were pooled together by replica ($n = 6$ per treatment) and mixed in 6 ml of LB medium. Each mixture of clones was injected into ten individual larvae, with time until death measured as a proxy for virulence. This procedure was performed in three independent repeats by injecting the same mixtures of bacterial clones into independent batches of larvae in separate experiments (total number of larvae = 420). To assess virulence of all evolved clones (Fig. 3c), infections were done independently using all of the individual PA14 clones from 3 d.p.i. from

the experiment shown in Extended Data Fig. 3 ($n$ = 1,008). Here (Fig. 3c), the bacterial inoculums were prepared individually for each clone by inoculating 200 µl LB medium with 5 µl bacterial sample from freezer stock, repeated for all individual clones in 96-well plates. Finally, to measure whether surface-based resistance against a LPS-specific phage was associated with similar virulence trade-offs (Extended Data Fig. 7), we isolated *P. aeruginosa* clones from six independent infection experiments with phage LMA2. A total of 10 clones per replicate experiment, isolated from 3 d.p.i., were phenotypically characterized to confirm resistance, and examined by PCR to exclude that resistance was CRISPR-based. All 10 clones with LPS-based resistance from the same replicate experiment were pooled together in 6 ml of LB medium ($n$ = 6), and infections of *G. mellonella* larvae were carried out as described above, with each mixture of clones injected into ten individual larvae, performed in three independent repeats (total number of larvae = 240). Before infection, all bacterial inoculums were grown overnight at 37 °C on an orbital shaker (180 r.p.m.) before being diluted by adding 20 µl to 180 µl of M9 salts. Cell density was then assayed by measuring optical density at 600 nm ($OD_{600}$), with 0.1 $OD_{600}$ being approximately $1 \times 10^8$ c.f.u. $ml^{-1}$, before being further diluted to approximately $1 \times 10^4$ c.f.u. $ml^{-1}$, which was subsequently used for infection by injecting 10 µl into the rear proleg of individual *G. mellonella* using a sterile syringe as previously described[22]. Optical density measurements and experimental repeats were taken into account during formal data analysis. After infection, larvae were incubated at 28 °C, with mortality monitored hourly for up to 48 h. For all independent experiments, a control experiment in which larvae were injected with just M9 salts was included. All work conforms to ethical regulations regarding the use of invertebrates, with approval from The University of Exeter ethics committee.

### Motility assays

Swarming motility of all evolved bacterial clones from the experiment shown in Extended Data Fig. 3c ($n$ = 1,008) was assayed by using a 96-well microplate pin replicator to stamp the individual clones on 1% agar before overnight growth at 37 °C. The diameters of the individual clones were then taken as a measure of motility (three replicas per clone).

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

All data used in this study are available on figshare at https://doi.org/10.6084/m9.figshare.9752903.


31. Martínez-García, E., Calles, B., Arévalo-Rodríguez, M. & de Lorenzo, V. pBAM1: an all-synthetic genetic tool for analysis and construction of complex bacterial phenotypes. *BMC Microbiol.* **11**, 38 (2011).
32. Goto, M. et al. Real-time PCR method for quantification of *Staphylococcus aureus* in milk. *J. Food Prot.* **70**, 90–96 (2007).
33. R Core Team. R: A Language and Environment for Statistical Computing. http://www.R-project.org/ (R Foundation for Statistical Computing, 2018).
34. Wickham, H. tidyverse: Easily Install and Load the 'Tidyverse'. R package version 1.2.1 https://cran.r-project.org/web/packages/tidyverse/index.html (2017).
35. Therneau, T. A Package for Survival Analysis in S. R package version 2.38 https://CRAN.R-project.org/package=survival (2015).



**Acknowledgements** We thank A. Buckling for critical reading of the manuscript, J. Common, E. Hesse and S. Meaden for comments on the manuscript, and J. P. Pirnay and D. de Vos for sharing clinical isolates of *S. aureus*, *A. baumannii* and *B. cenocepacia*. This work was supported by grants from the ERC (ERC-STG-2016-714478 - EVOIMMECH) and the NERC (NE/M018350/1), which were awarded to E.R.W.


**Author contributions** Conceptualization of the study was done by E.O.A. and E.R.W. Experimental design was carried out by E.O.A., A.M.L., C.R. and E.R.W. Adsorption and infection assays were done by E.O.A. All evolution experiments were performed by E.O.A., E.P. and I.M. E.O.A. performed the DNA extractions and qPCR reactions, and the competition experiments, virulence assays and motility assays were performed by E.O.A. and E.P. Formal analysis of results was done by E.O.A., E.P., C.R. and E.R.W. The original draft was written by E.O.A., with later edits and reviews by E.O.A. and E.R.W.

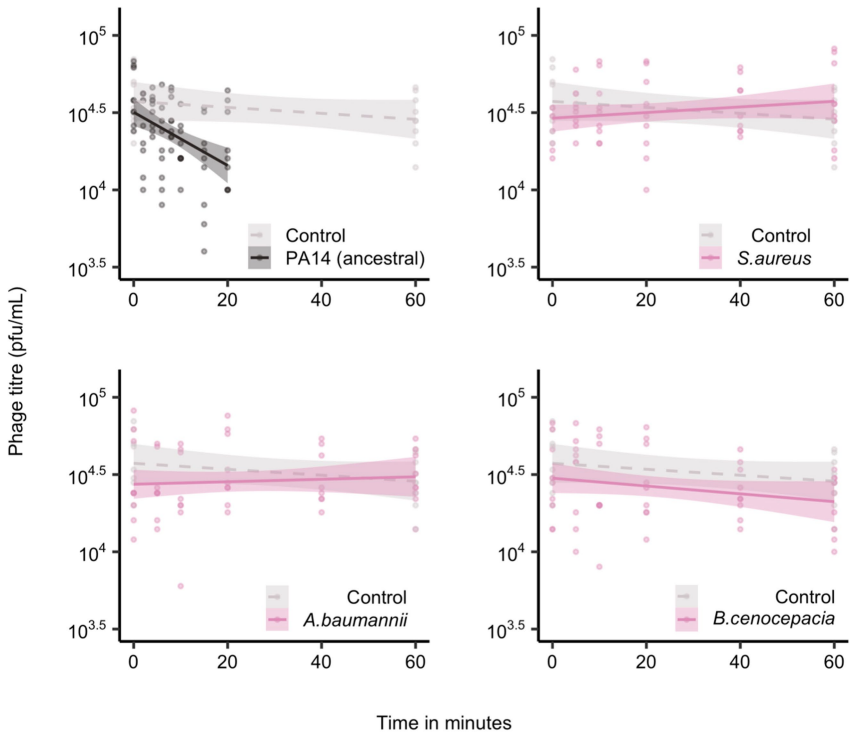
**Competing interests** The authors declare no competing interests.

**Additional information**

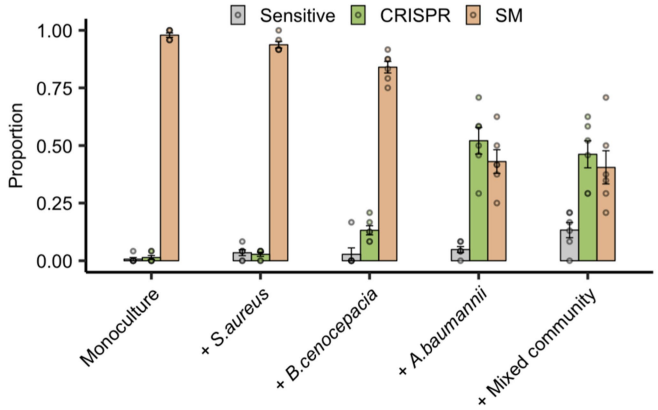

**Supplementary information** is available for this paper at https://doi.org/10.1038/s41586-019-1662-9.



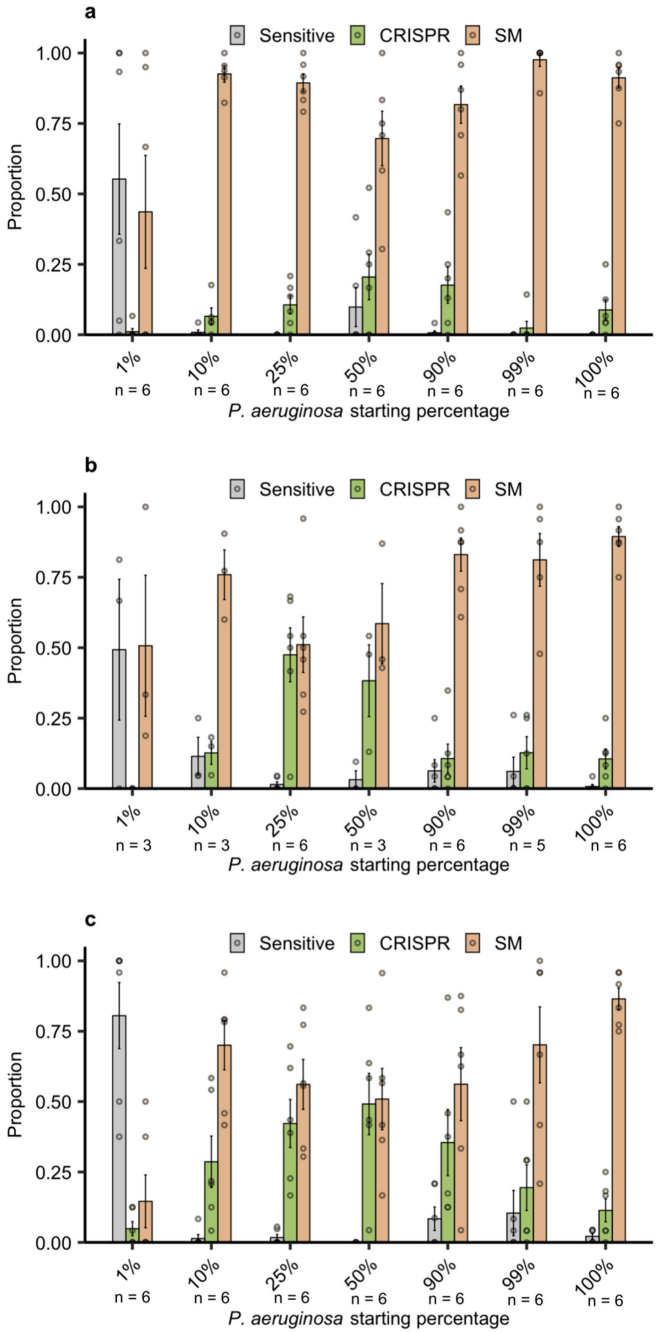
**Correspondence and requests for materials** should be addressed to E.O.A. or E.R.W.




**Extended Data Fig. 1 | Only *P. aeruginosa* adsorbs phage DMS3vir.** Phage levels (in p.f.u. $ml^{-1}$) in minutes after infection of *P. aeruginosa* PA14 and three other bacterial species ($n = 84$ biologically independent replicates). Controls were carried out in the absence of bacteria. Here, the lines are regression slopes with shaded areas corresponding to 95% confidence intervals. Linear model: effect of *P. aeruginosa* on phage titre over time; $t = -3.37$, $P = 0.0009$; *S. aureus*; $t = 1.63$, $P = 0.11$; *A. baumannii*; $t = 1.20$, $P = 0.23$; *B. cenocepacia*; $t = -0.27$, $P = 0.79$; overall model fit; $F_{9,235} = 4.33$, adjusted $R^2 = 0.11$, $P = 3.17 \times 10^{-5}$.

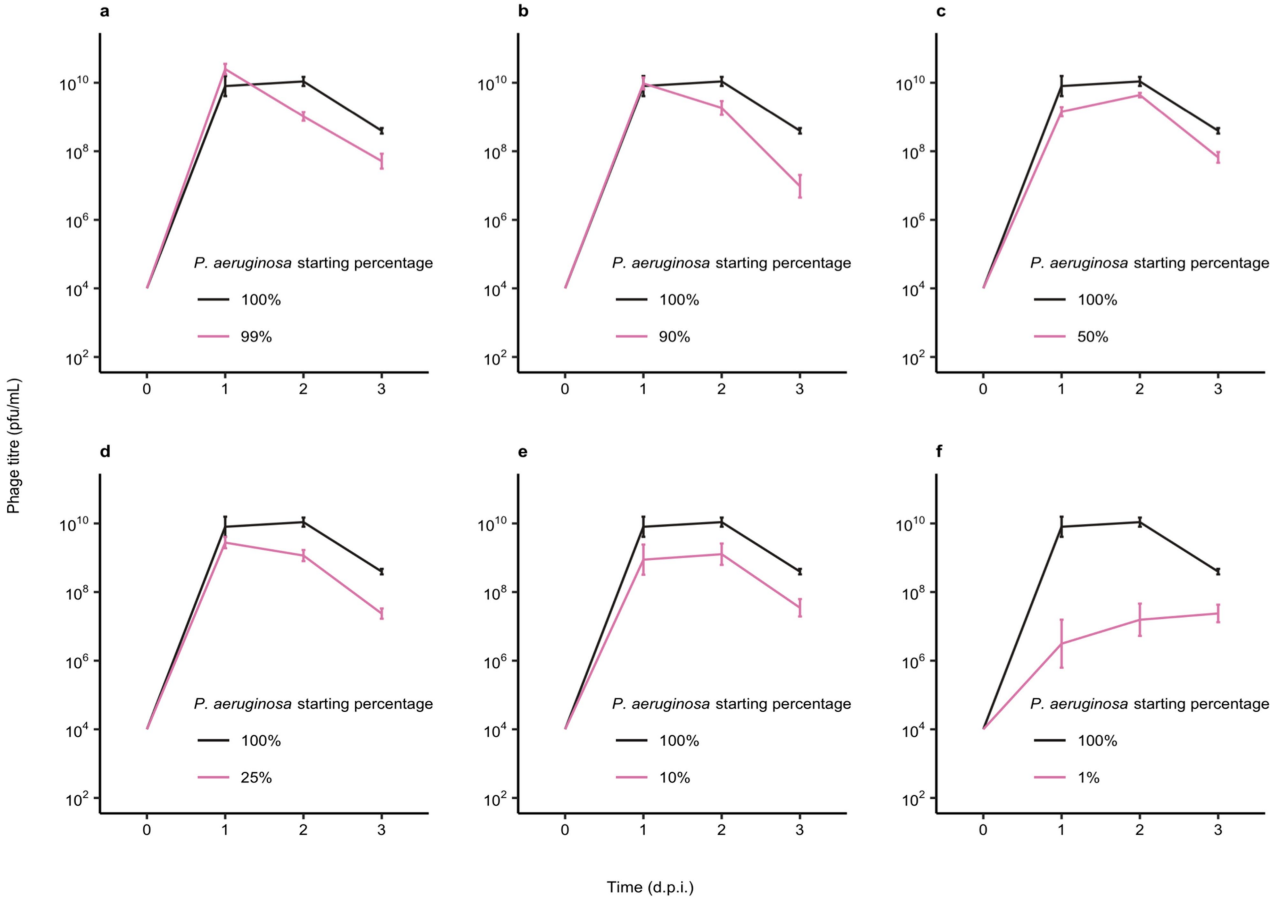



**Extended Data Fig. 2 | Enhanced CRISPR resistance evolution in ASM.** Proportion of *P. aeruginosa* that acquired surface modification or CRISPR-based immunity (or remained sensitive) 3 d.p.i. with phage DMS3vir when grown in ASM (6 replicates per treatment, with 24 colonies screened from each replicate, $n = 30$ biologically independent replicates). Deviance test: relationship between community composition and CRISPR; residual deviance (25, $n = 30$) = 1.26, $P = 2.2 \times 10^{-16}$; Tukey contrasts: monoculture versus mixed; $z = -5.30$, $P = 1 \times 10^{-4}$; monoculture versus *A. baumannii*; $z = -5.60$, $P = 1 \times 10^{-4}$; monoculture versus *B. cenocepacia*; $z = -2.80$, $P = 0.02$; monoculture versus *S. aureus*; $z = -0.76$, $P = 0.93$. Data are mean ± s.e.m.

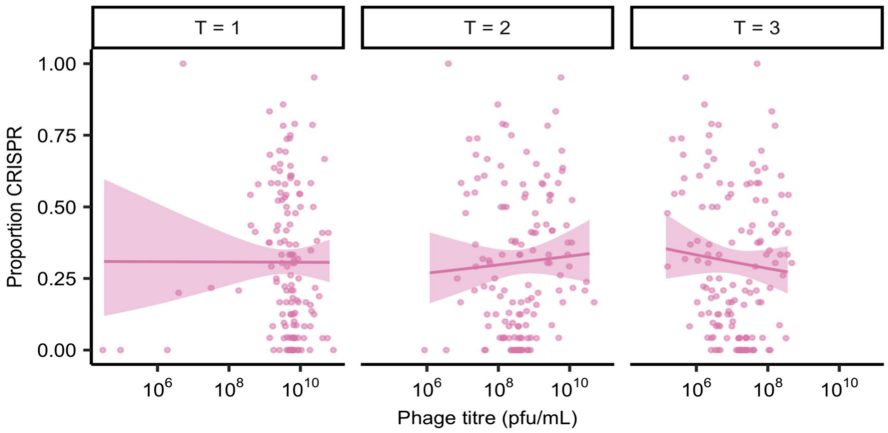



**Extended Data Fig. 3** | See next page for caption.

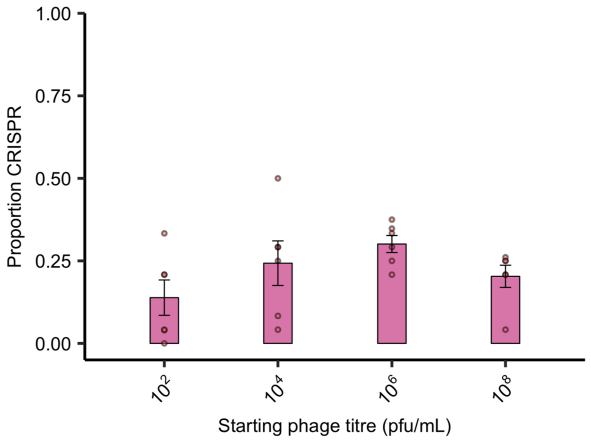
**Extended Data Fig. 3 | Increased evolution of CRISPR-based resistance across a range of microbial community compositions over time.** Proportion of *P. aeruginosa* that acquired surface modification or CRISPR-based immunity (or remained sensitive) at up to 3 d.p.i. with phage DMS3vir when grown either in monoculture (100%) or in polyculture mixtures consisting of the mixed microbial community but with varying starting percentages of *P. aeruginosa* based on volume (6 replicates for most samples, with 24 colonies per replicate, $n = 42$ biologically independent replicates for **a**, $n = 32$ biologically independent replicates for **b**, and $n = 42$ biologically independent replicates for **c**). **a**, Resistance evolution at 1 d.p.i. Data are mean ± s.e.m. Deviance test: relationship between CRISPR and *P. aeruginosa* starting percentage at time point 1; residual deviance ($35, n = 42$) = 4.42, $P = 0.004$; 1%; $z = -3.27$, $P = 0.002$; 10%; $z = 1.21$, $P = 0.23$; 25%; $z = 1.62$, $P = 0.11$; 50%; $z = 2.20$, $P = 0.034$; 90%; $z = 2.07$, $P = 0.046$; 99%; $z = 0.47$, $P = 0.65$; 100%; $z = 1.47$, $P = 0.15$. **b**, Resistance evolution at 2 d.p.i. Data are mean ± s.e.m. Deviance test: relationship between CRISPR and *P. aeruginosa* starting percentage at time point 2; residual deviance ($25, n = 32$) = 3.86, $P = 2.51 \times 10^{-6}$; 1%; $z = -2.14$, $P = 0.04$; 10%; $z = 1.19$, $P = 0.25$; 25%; $z = 2.07$, $P = 0.049$; 50%; $z = 1.89$, $P = 0.07$; 90%; $z = 1.12$, $P = 0.27$; 99%; $z = 1.21$, $P = 0.24$; 100%; $z = 1.11$, $P = 0.28$. **c**, Resistance evolution at 3 d.p.i. Data are mean ± s.e.m. Deviance test: relationship between CRISPR and *P. aeruginosa* starting percentage at time point 3; residual deviance ($35, n = 42$) = 8.24, $P = 0.0004$; 1%; $z = -3.38$, $P = 0.002$; 10%; $z = 2.12$, $P = 0.04$; 25%; $z = 2.77$, $P = 0.009$; 50%; $z = 3.07$, $P = 0.004$; 90%; $z = 2.46$, $P = 0.019$; 99%; $z = 1.55$, $P = 0.13$; 100%; $z = 0.87$, $P = 0.39$.



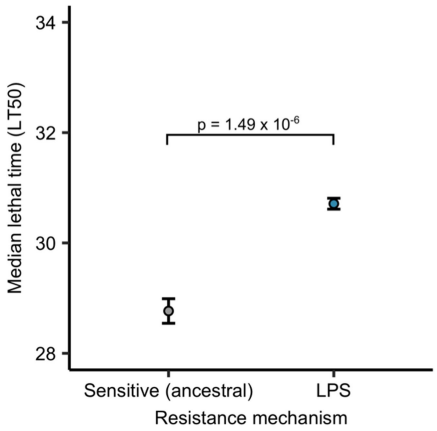

**Extended Data Fig. 4 | Microbial community composition affects phage epidemic size.** The DMS3vir phage titres (in p.f.u. $ml^{-1}$) over time up to 3 d.p.i. of *P. aeruginosa* grown either in monoculture (100%) or in polyculture mixtures as shown in Extended Data Fig. 3. Each data point represents the mean, error bars denote s.e.m. ($n = 171$ independent biological samples). Two-way ANOVA: overall effect of *P. aeruginosa* starting percentage on phage titre; $F_{6,105} = 14.84$, $P = 1.1 \times 10^{-12}$.




**Extended Data Fig. 5 | No correlation between phage epidemic size and evolution of CRISPR resistance.** The correlation between the proportion of evolved phage-resistant clones with CRISPR-based resistance and the phage epidemic sizes (in p.f.u. $ml^{-1}$) in the presence of other bacterial species, using data taken from experiments shown in Fig. 1, Extended Data Figs. 2, 3c and 6 ($n = 137$ biologically independent samples per time point). Correlations are separated by day, as phage titres were measured daily. Here, the lines are regression slopes, with shaded areas corresponding to 95% confidence intervals. Pearson's product–moment correlation tests between phage titres (at each day after infection) and levels of CRISPR-based resistance: $T = 1$; $t_{136} = -0.02$, $P = 0.98$, $R^2 = -0.002$; $T = 2$; $t_{136} = 0.59$, $P = 0.55$, $R^2 = 0.05$; $T = 3$; $t_{136} = -0.90$, $P = 0.37$, $R^2 = -0.08$.




**Extended Data Fig. 6 | Starting phage titre does not affect CRISPR evolution in the presence of a microbial community.** Proportion of *P. aeruginosa* that acquired CRISPR-based resistance at 3 d.p.i. with varying starting titres of phage DMS3vir when grown in polyculture (6 replicates per treatment, with 24 colonies per replicate, $n = 24$ biologically independent replicates). Deviance test: start phage and CRISPR; residual deviance (20, $n = 24$) = 2.00, $P = 0.13$; Tukey contrasts: $10^2$ versus $10^4$; $z = -1.52$, $P = 0.42$; $10^4$ versus $10^6$; $z = -0.76$, $P = 0.87$; $10^6$ versus $10^8$; $z = 1.31$, $P = 0.56$; $10^2$ versus $10^6$; $z = -2.24$, $P = 0.11$; $10^2$ versus $10^8$; $z = -0.99$, $P = 0.75$; $10^4$ versus $10^8$; $z = 0.56$, $P = 0.94$. Data are mean ± s.e.m.




**Extended Data Fig. 7 | LPS-based phage resistance also affects in vivo virulence.** Time until death (given as median ± one standard error) for *G. mellonella* larvae infected with PA14 clones that evolved phage resistance by LPS modification, compared to the phage-sensitive ancestral ($n = 209$ biologically independent samples). Cox proportional hazards model with Tukey contrasts: sensitive (ancestral) versus LPS; $z = 4.81$, $P = 1.49 \times 10^{-6}$. overall model fit; $LRT_3 = 44.94$, $P = 1 \times 10^{-9}$.

nature research

Corresponding author(s): Edze Westra
Ellinor Alseth

Last updated by author(s): Sep 2, 2019

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☒ | ☐ | A description of all covariates tested |
| ☒ | ☐ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ | For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☒ | ☐ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒ | ☐ | Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |

*Our web collection on statistics for biologists contains articles on many of the points above.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | No software was used for data collection |
| Data analysis | All data analyses were done using R version 3.5.1., and the Tidyverse package version 1.2.1.. The survival analyses were done using the Survival package version 2.38. |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data used in this study is available on figshare at 10.6084/m9.figshare.9752903 (will be made public prior to publication).

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Sample size | All in vitro experiments were performed in 6 independent replicates for sufficient statistical power, in line with common practice in experimental evolution studies. Larvae infections were performed in 10 replicates with 3 independent repeats for sufficient statistical power, and in line with common practice in Galleria virulence assays. All sample sizes were determined based on common practice in experimental evolution studies and infection assays using Galleria. |
| Data exclusions | No data were excluded from the analyses |
| Replication | We used 6 independent biological replicates per treatment. For virulence assays, there were 10 larvae per treatment, with 3 independent repeats. All observations were reproducible. |
| Randomization | Not relevant to our study - single experimental manipulations of individual variables. |
| Blinding | Blinding not relevant - single experimental manipulations of individual variables. |

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials & experimental systems

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ Antibodies |
| ☒ | ☐ Eukaryotic cell lines |
| ☒ | ☐ Palaeontology |
| ☐ | ☒ Animals and other organisms |
| ☒ | ☐ Human research participants |
| ☒ | ☐ Clinical data |

## Methods

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ ChIP-seq |
| ☒ | ☐ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |

## Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

| | |
|---|---|
| Laboratory animals | We used Galleria mellonella (wax moth) larvae. |
| Wild animals | The study did not involve wild animals |
| Field-collected samples | The study did not involve samples collected from the field |
| Ethics oversight | Ethics approval for the use of invertebrates was given by The University of Exeter ethics committee. |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Article

# Heterogeneity in old fibroblasts is linked to variability in reprogramming and wound healing

https://doi.org/10.1038/s41586-019-1658-5

Received: 31 March 2016

Accepted: 5 September 2019

Published online: 23 October 2019

Salah Mahmoudi[1,11], Elena Mancini[1,11], Lucy Xu[1,2], Alessandra Moore[3,4], Fereshteh Jahanbani[1], Katja Hebestreit[1], Rajini Srinivasan[4,5], Xiyan Li[1], Keerthana Devarajan[1], Laurie Prélot[1], Cheen Euong Ang[4,6,7], Yohei Shibuya[4,7], Bérénice A. Benayoun[1,10], Anne Lynn S. Chang[8], Marius Wernig[4,7], Joanna Wysocka[4,5], Michael T. Longaker[3,4], Michael P. Snyder[1] & Anne Brunet[1,9]*

Age-associated chronic inflammation (inflammageing) is a central hallmark of ageing[1], but its influence on specific cells remains largely unknown. Fibroblasts are present in most tissues and contribute to wound healing[2,3]. They are also the most widely used cell type for reprogramming to induced pluripotent stem (iPS) cells, a process that has implications for regenerative medicine and rejuvenation strategies[4]. Here we show that fibroblast cultures from old mice secrete inflammatory cytokines and exhibit increased variability in the efficiency of iPS cell reprogramming between mice. Variability between individuals is emerging as a feature of old age[5–8], but the underlying mechanisms remain unknown. To identify drivers of this variability, we performed multi-omics profiling of fibroblast cultures from young and old mice that have different reprogramming efficiencies. This approach revealed that fibroblast cultures from old mice contain 'activated fibroblasts' that secrete inflammatory cytokines, and that the proportion of activated fibroblasts in a culture correlates with the reprogramming efficiency of that culture. Experiments in which conditioned medium was swapped between cultures showed that extrinsic factors secreted by activated fibroblasts underlie part of the variability between mice in reprogramming efficiency, and we have identified inflammatory cytokines, including TNF, as key contributors. Notably, old mice also exhibited variability in wound healing rate in vivo. Single-cell RNA-sequencing analysis identified distinct subpopulations of fibroblasts with different cytokine expression and signalling in the wounds of old mice with slow versus fast healing rates. Hence, a shift in fibroblast composition, and the ratio of inflammatory cytokines that they secrete, may drive the variability between mice in reprogramming in vitro and influence wound healing rate in vivo. This variability may reflect distinct stochastic ageing trajectories between individuals, and could help in developing personalized strategies to improve iPS cell generation and wound healing in elderly individuals.

Several studies have investigated the effect of ageing and senescence on reprogramming[9–12], but a systematic evaluation of how ageing influences reprogramming is lacking. We examined the influence of old age on the inflammatory profile of fibroblasts and their ability to reprogram to iPS cells (Fig. 1a). Using cytokine profiling, we compared the systemic milieu (plasma) and conditioned medium from primary fibroblast cultures from young (3 months) and old (28–29 months) mice (Fig. 1a). Plasma from old mice showed increased levels of pro-inflammatory cytokines (for example, IL-6 and TNF), anti-inflammatory cytokines (for example, IL-4), and chemokines and growth factors (for example, CSF1 (also known as MCSF)) compared to plasma from young mice (Fig. 1b, Extended Data Fig. 1a, b and Supplementary Table 1a). Conditioned medium from primary fibroblast cultures from the ears of old mice also showed enhanced levels of pro- and anti-inflammatory cytokines (for example, IL-6 and TNF, and IL-4, respectively; (Fig. 1b, Extended Data Fig. 1c, d and Supplementary Table 1b). Similarly, inflammatory cytokines increased with

[1]Department of Genetics, Stanford University, Stanford, CA, USA. [2]Department of Biology, Stanford University, Stanford, CA, USA. [3]Department of Surgery, Division of Plastic and Reconstructive Surgery, Stanford University, Stanford, CA, USA. [4]Institute for Stem Cell Biology and Regenerative Medicine, Stanford University, Stanford, CA, USA. [5]Department of Chemical and Systems Biology, Stanford University, Stanford, CA, USA. [6]Department of Bioengineering, Stanford University, Stanford, CA, USA. [7]Department of Pathology, Stanford University, Stanford, CA, USA. [8]Department of Dermatology, Stanford University, Stanford, CA, USA. [9]Glenn Laboratories for the Biology of Aging, Stanford University, Stanford, CA, USA. [10]Present address: Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA. [11]These authors contributed equally: Salah Mahmoudi, Elena Mancini. *e-mail: anne.brunet@stanford.edu

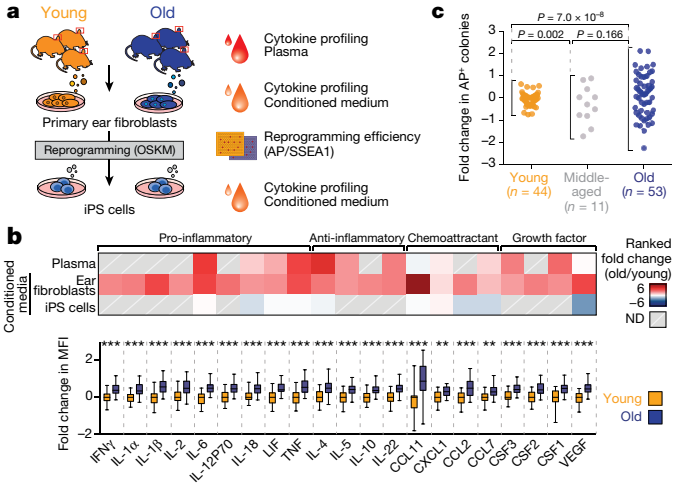



**Fig. 1 | Primary fibroblasts from old mice secrete inflammatory cytokines and show increased variability in reprogramming efficiency between mice.** **a**, Experimental schematic. Young mice, 3 months old; old mice, 28–29 months old. OSKM, *OCT4*, *SOX2*, *KLF4* and *MYC*. **b**, Top, age-dependent changes in cytokine levels in plasma and conditioned medium from fibroblasts or iPS cells (Extended Data Fig. 1a, g, h). ND, not detected. Bottom, cytokine profiles of conditioned medium from primary cultures (passage 3) of ear fibroblasts from young (3 months, $n = 24$) and old (29 months, $n = 24$) male mice (3 independent experiments). Box-and-whisker plots of $\log_2$-transformed fold change in mean fluorescence intensity (MFI) compared to the median of young fibroblasts. Box plots depict median and interquartile range, with whiskers indicating minimum and maximum values. $^{**}P < 0.01$, $^{***}P < 0.001$; two-tailed Wilcoxon rank-sum test with Benjamini–Hochberg correction. Exact *P* values can be found in Supplementary Table 1b. **c**, Reprogramming efficiency assessed by alkaline phosphatase (AP) staining of cultures of ear fibroblasts obtained from young (3 months, $n = 44$), middle-aged (12 months, $n = 11$) and old (28–29 months, $n = 53$) mice (7 independent experiments). The $\log_2$-transformed fold change over the median of young mice is shown. Each dot represents a fibroblast culture from one mouse. *P* values, Fligner–Killeen test to assess differences in variance between age groups with Benjamini–Hochberg correction.

age in conditioned medium from lung fibroblasts and human primary fibroblasts (Extended Data Fig. 1e, f and Supplementary Table 1c, d). Thus, primary cultures of fibroblasts from old mice exhibit a secretory inflammatory profile that overlaps in part with that of the systemic milieu (Fig. 1b and Extended Data Fig. 1h).

To systematically test the effect of age on iPS cell reprogramming, we derived independent fibroblast cultures from a total of 108 young, middle-aged and old mice. We induced reprogramming by expressing human *OCT4* (also known as *POU5F1*), *KLF4*, *SOX2* and *MYC*[13], and assessed reprogramming efficiency using alkaline phosphatase (AP) and stage-specific embryonic antigen 1 (SSEA1) staining[10] (Fig. 1a and Extended Data Fig. 1i–l). We did not observe a significant change in mean reprogramming efficiency with age (Fig. 1c and Extended Data Fig. 1l). However, there was increased variability between mice in reprogramming efficiency with age, with cultures from some old mice reprogramming better and some worse than cultures from young mice (Fig. 1c and Extended Data Fig. 1l). A similar age-dependent increase in variability in reprogramming efficiency was observed in chest fibroblast cultures (Extended Data Fig. 1m). Reprogramming efficiency appeared to be inherent to each culture (derived from an individual mouse), as the same culture exhibited largely consistent reprogramming efficiency to iPS cells between independent experiments or to induced neurons (Extended Data Fig. 1n, o). This increased variability in reprogramming efficiency between fibroblast cultures from different old mice could reflect distinct stochastic ageing trajectories.

Variability between old individuals has been observed for several biological features[5–8]. However, most studies were performed in humans, in

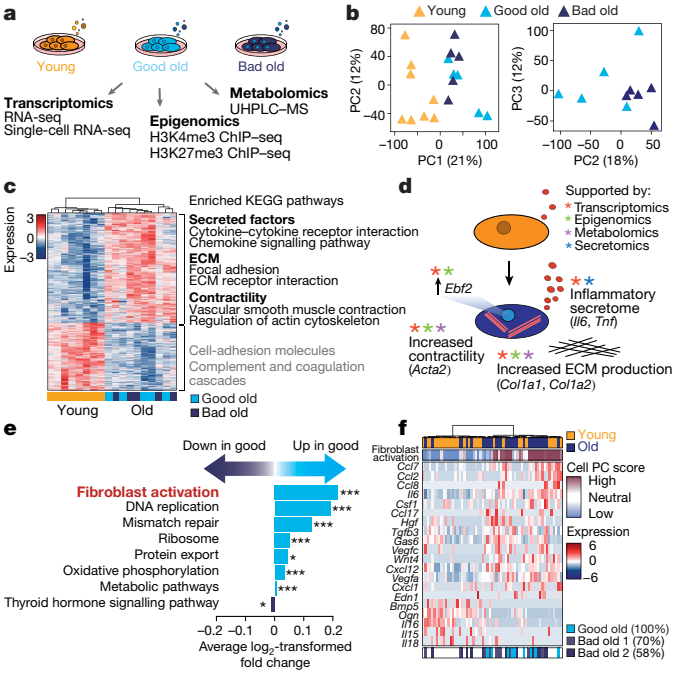



**Fig. 2 | Old fibroblast cultures exhibit a signature of an inflammatory activated state, which is associated with variability in reprogramming efficiency.** **a**, Multi-omics characterization of fibroblast cultures. ChIP–seq, chromatin immunoprecipitation followed by sequencing; UHPLC–MS, ultra-high performance liquid chromatography–tandem mass spectrometry. **b**, Principal component (PC) analysis of transcriptomes cultures of ear fibroblasts from young (3 months, $n = 8$) and old (29 months, $n = 10$) (left) or only old (right) mice (3 independent experiments). Old cultures were either good (high reprogramming efficiency) or bad (low reprogramming efficiency) (Supplementary Table 2a). **c**, Heat map of significantly differentially expressed genes (determined by DESeq2) between young and old fibroblasts described in **b** and enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. Pathways are colour coded according to significance (one-sided Fisher's exact test with Benjamini–Hochberg correction; black, false-discovery rate (FDR)-adjusted $P < 0.05$; grey, FDR-adjusted $P < 0.15$; Supplementary Table 2b, c). ECM, extracellular matrix. **d**, Summary of the multi-omics profiling of young and old fibroblast cultures (Extended Data Fig. 2). **e**, Pathway enrichment analysis of KEGG pathways associated with good or bad reprogramming efficiency. Overlapping significant (FDR-adjusted $P < 0.05$) KEGG pathways identified in a regression analysis from bad to good reprogramming efficiencies ($n = 18$) and in a separate analysis comparing the five highest and five lowest reprogramming efficiency cultures (exact *P* values can be found in Supplementary Table 3b, d). $^{*}P < 0.05$, $^{***}P < 0.001$; two-sided nominal *P* value with Benjamini–Hochberg correction. **f**, PAGODA of single-cell RNA-seq from young ($n = 30$ cells), good and bad old ($n = 31$ cells) fibroblasts. Top heat maps, PAGODA clustering of cells. For cell PC score, maroon and blue colours indicate increased and decreased expression of the associated gene sets, respectively. Middle heat map, expression of specific cytokine genes (Extended Data Fig. 4g). Bottom heat map, single cells from good and bad old fibroblast cultures. Gene expression is shown as VST-transformed (variance stabilizing transformation, implemented in DESeq2) read counts scaled row-wise.

which genetic and environmental differences also have a role. We used the controlled mouse system to understand the stochastic variability in reprogramming efficiency between cultures from old mice. Using a multi-omics approach, we profiled the transcriptomes, epigenomes and metabolomes of young fibroblasts as well as old fibroblasts that reprogrammed well (good old) or poorly (bad old) (Fig. 2a and Supplementary Table 2a). Principal component analysis and unsupervised hierarchical clustering showed a separation between young and old fibroblasts across datasets (Fig. 2b and Extended Data Fig. 2a–h). Principal component analysis also revealed some separation between the

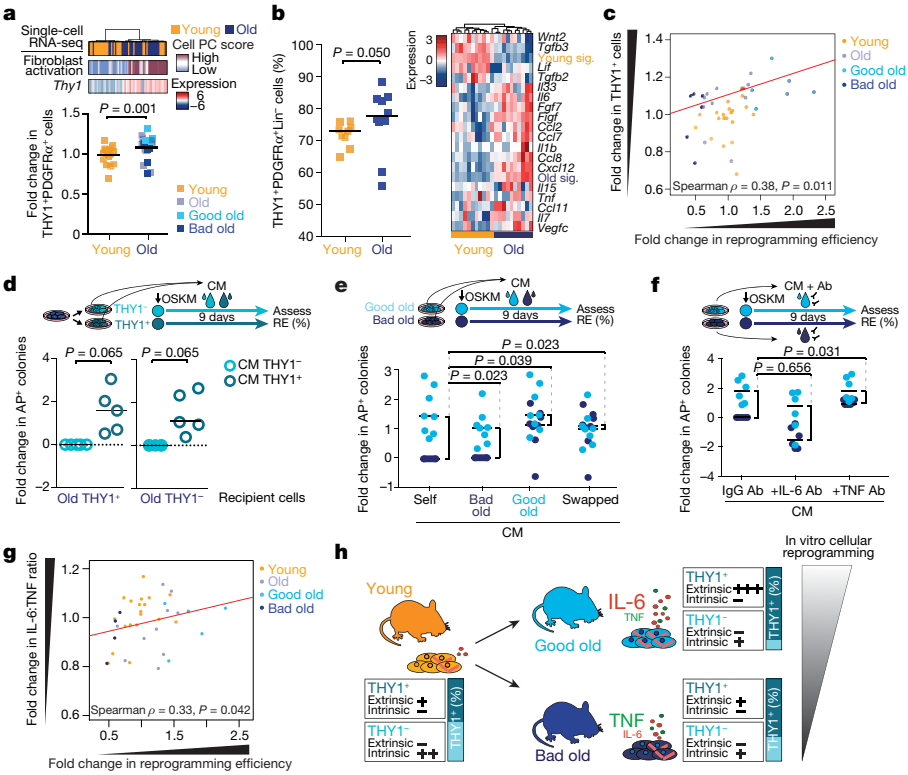



**Fig. 3 | Age-associated increase in activated fibroblasts and the cytokines that they secrete drive part of the variability in reprogramming between mice. a**, Top, PAGODA clustering of single-cell RNA-seq data from young and old fibroblasts as in Fig. 2f, showing *Thy1* expression. Bottom, proportion of THY1$^+$PDGFRα$^+$ fibroblasts in fibroblast cultures of young (3 months, $n = 21$) and old (29 months, $n = 23$) mice measured by FACS (3 independent experiments). Fold changes were calculated relative to the median of young mice. *P* value, two-tailed Wilcoxon rank-sum test. Each dot represents a culture from one mouse. Lines depict median. **b**, Left, Percentage of THY1$^+$PDGFRα$^+$Lin$^-$ out of all PDGFRα$^+$Lin$^-$ fibroblasts isolated from ears of young mice (3–4 months, $n = 9$ replicates, each with 2–3 mice) and old mice (24–26 months, $n = 10$ replicates, each with 2–3 mice) analysed by FACS (3 independent experiments). *P* value as in **a**. Each dot represents a replicate, with cells pooled from 2–3 mice. Lines depict median. Right, heat map of the expression of specific cytokine genes from population RNA-seq of fibroblasts. VST-transformed read counts are shown scaled row-wise. Young sig. and old sig. indicate the average expression of genes that are significantly downregulated and upregulated with age, respectively. **c**, Spearman's correlation between the proportion of THY1$^+$PDGFRα$^+$ (THY1$^+$) fibroblasts in a given culture (quantified by FACS) and the reprogramming efficiency (assessed as in Fig. 1c) of that culture (ages as in **a**; young, $n = 21$; old, $n = 23$; 3 independent experiments). Fold changes relative to the median of young mice. *P* values, two-sided algorithm AS 89 in R. Each dot represents a culture from one mouse. **d**, Reprogramming efficiency (RE) of FACS-sorted old THY1$^+$PDGFRα$^+$ (THY1$^+$) and THY1$^-$PDGFRα$^+$ (THY1$^-$) fibroblasts treated daily with conditioned medium (CM) from THY1$^-$PDGFRα$^+$ or THY1$^+$PDGFRα$^+$ fibroblasts from the same original culture. $\log_2$-transformed fold change relative to THY1$^-$PDGFRα$^+$ fibroblasts treated with conditioned medium from THY1$^-$PDGFRα$^+$ fibroblasts ($n = 5$ old mice, 4 independent experiments). *P* values, two-tailed Wilcoxon signed-rank test. Each dot represents a culture from one mouse. Lines depict median. **e**, Reprogramming efficiency of pairs of good old and bad old fibroblast cultures treated with their own conditioned medium (self conditioned medium) or conditioned medium from the other group (swapped conditioned medium). $\log_2$-transformed fold change relative to bad old self conditioned medium. $n = 8$ pairs of good and old bad cultures (5 independent experiments). *P* values, two-tailed Wilcoxon signed-rank test with Benjamini–Hochberg correction. Each dot represents a culture from one mouse. Lines depict median. **f**, Reprogramming efficiency of pairs of good old and bad old fibroblast cultures treated with their own conditioned medium, which was pretreated with blocking antibodies. $\log_2$-transformed fold change in reprogramming efficiency relative to bad old conditioned medium treated with IgG antibodies. $n = 6$ pairs of good old and bad old cultures (4 independent experiments). *P* values, two-tailed Wilcoxon signed-rank test with Benjamini–Hochberg correction. Each dot represents a culture from one mouse. Lines depict median. **g**, Spearman's correlation between conditioned medium and the ratio of IL-6 and TNF levels in the conditioned medium (young, $n = 19$; old, $n = 18$; ages as in **a**; 2 independent experiments). Fold change relative to the median of young mice. *P* values, two-sided algorithm AS 89 in R. Each dot represents a culture from one mouse. **h**, Model for the increased variability in cellular reprogramming between mice in vitro.

transcriptomes and metabolomes of good old and bad old cultures (Fig. 2b and Extended Data Fig. 2i, j).

Old fibroblasts showed transcriptional enrichment of pathways related to secreted factors (for example, cytokine signalling), extracellular matrix, contractility, inflammation and wound healing (Fig. 2c, d, Extended Data Fig. 2k, l and Supplementary Table 2b–e). These features are characteristic of activated fibroblasts (also known as myofibroblasts), which are normally involved in tissue repair[2,3,14,15]. Indeed, the 'fibroblast activation' gene set was enriched in the old fibroblast transcriptomes (Extended Data Fig. 2m and Supplementary Table 2f). Epigenomic and metabolomics changes supported this fibroblast activation signature (Fig. 2d, Extended Data Fig. 2n–t and Supplementary Table 2g–m). The transcription factor EBF2, which shows increased expression in old fibroblasts, was identified as a potential driver of this activated fibroblast signature (Fig. 2d, Extended Data Fig. 2q, u and Supplementary Table 2n). Primary fibroblast cultures from elderly humans also exhibited increased *EBF2* and cytokine-related pathway expression (Extended Data Fig. 2v, Supplementary Table 2o, p). Notably, fibroblast activation was a top feature associated with good reprogramming of old fibroblasts in both transcriptomic and epigenomic datasets (Fig. 2e, Extended Data Fig. 2w and Supplementary Table 3a–f). Hence, the fibroblast activation signature is enriched in old fibroblasts and correlates with the variability between mice in reprogramming.

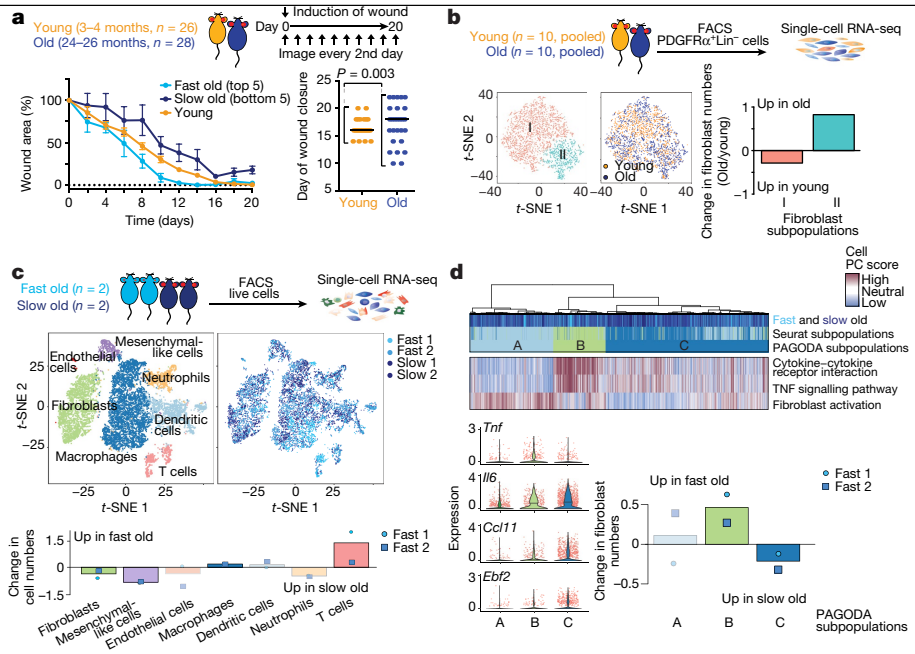



**Fig. 4 | Wound healing rate is variable between old mice and correlates with fibroblast subpopulations with distinct cytokine signatures. a**, Ear wound healing assays in young (3–4 months, $n = 26$) and old (24–26 months, $n = 28$) mice (2 independent experiments). Left, ear wound healing curves from young mice and the five fastest- and five slowest-healing old mice. Percentage of wound area that remains on the indicated day (mean ± s.d.) Right, day of ear wound closure in young and old mice. Each dot represents one mouse. Line marks median. *P* values, Fligner–Killeen test to assess difference in variance between age groups. **b**, Single-cell RNA-seq of FACS-sorted PDGFRα$^+$Lin$^-$ cells from the ear wounds of young mice (3–4 months, cells pooled from $n = 10$ mice) or old mice (24–26 months, cells pooled from $n = 10$ mice), 7 days after induction of wounds. Left, *t*-distributed stochastic neighbour embedding (*t*-SNE) clustering of cells (3,036 total; 1,592 young, 1,444 old) coloured by Seurat clusters or age. Right, $\log_2$-transformed fold change in the subpopulations between wounds of young and old mice. **c**, Single-cell RNA-seq of live cells from entire wounds of old mice (24 months) with fast-healing ($n = 2$) and slow-healing ($n = 2$) trajectories, 7 days after induction of wounds. *t*-SNE clustering of cells ($n = 10{,}797$ total), coloured by Seurat clusters or mouse (slow old 1, $n = 3{,}761$; slow old 2, $n = 2{,}127$; fast old 1, $n = 2{,}533$; fast old 2, $n = 2{,}376$). Bottom, $\log_2$-transformed fold change in the cell types between wounds from fast-healing compared to slow-healing old mice. **d**, PAGODA clustering of cells ($n = 2{,}678$ total; slow old 1, $n = 1{,}087$; slow old 2, $n = 551$; fast old 1, $n = 441$; fast old 2, $n = 599$) identified as fibroblasts in **c**. Top heat map, single cells from wounds from old mice with fast- and slow-healing trajectories. Bottom heat map, separation of cells based on principal component scores for a subset of the top significantly overdispersed gene sets. For cell PC score, maroon and blue colours indicate generally increased and decreased expression of the associated gene sets, respectively. $\log_2$-transformed and normalized gene expression values calculated by Seurat and scaled row-wise. Bottom left, $\log_2$-normalized expression values of relevant genes. Each dot represents a single cell. Line marks median. Bottom right, $\log_2$-tranformed fold change in the number of cells in each of the three fibroblast subpopulations identified by PAGODA.

We wondered whether age-dependent cellular heterogeneity[8,16–19] could contribute to the variability between individual mice. To determine whether fibroblast cultures are heterogeneous, we performed single-cell RNA sequencing (RNA-seq) on young, good old and bad old cultures. Although the number of single cells profiled was low, the good old culture contained a higher proportion of activated cells compared to the two bad old cultures (Fig. 2f, Extended Data Fig. 4a–g and Supplementary Table 3g). Thus, the proportion of activated fibroblasts may be linked to the variability in reprogramming between individual cultures.

We validated that old fibroblast cultures were enriched in activated cells by staining for α-smooth muscle actin (αSMA), a marker of activated fibroblasts[2,3,14,15] (Extended Data Fig. 5a). These activated fibroblasts were proliferating and did not exhibit senescence markers (for example, *p16$^{Ink4a}$*) (Extended Data Fig. 5b–e). Fluorescence-activated cell sorting (FACS) analysis of the pan-fibroblast marker PDGFRα[3,14,19] as well as THY1[20], which correlates with the activated fibroblast signature, confirmed that old fibroblast cultures contained higher proportions of THY1$^+$PDGFRα$^+$ cells (Fig. 3a and Supplementary Table 4a–c). THY1$^+$PDGFRα$^+$ cells expressed fibroblast activation markers, inflammatory cytokines and *Ebf2* (Extended Data Fig. 5f). *Ebf2* knockdown in these cells reduced expression of fibroblast activation genes (for example, *Acta2* (which encodes αSMA), *Il6* and *Ccl11* (also known as *Eotaxin*)), whereas *Ebf2* overexpression in young fibroblasts induced expression of cytokines (for example, *Il6*; Extended Data Fig. 5g, h). In vivo FACS analysis also revealed a higher proportion of THY1$^+$PDGFRα$^+$ fibroblasts in the ears of old mice (Fig. 3b), and these fibroblasts exhibited a fibroblast activation signature with expression of inflammatory cytokines (Fig. 3b, Extended Data Fig. 5i–k and Supplementary Table 4d–g). Thus, activated fibroblasts are enriched in old cultures and old tissues in vivo.

Notably, FACS analysis of fibroblast cultures corroborated the positive correlation between the proportion of activated (THY1$^+$PDGFRα$^+$) fibroblasts in a culture and the ability of this culture to reprogram (Fig. 3c and Extended Data Fig. 5l–n). Reprogramming efficiency also correlated positively with proliferation and negatively with senescence (Extended Data Fig. 5o, p). Thus, the proportion of activated fibroblasts, though not more variable with age, correlates positively with reprogramming efficiency.

We next investigated how activated fibroblasts influence reprogramming efficiency. Activated THY1$^+$PDGFRα$^+$ fibroblasts intrinsically reprogrammed less efficiently than their non-activated THY1$^-$PDGFRα$^+$ counterparts (Extended Data Fig. 5q, r). By contrast, conditioned medium from activated fibroblasts enhanced reprogramming (of both activated and non-activated fibroblasts) compared to medium from non-activated fibroblasts (Fig. 3d, Extended Data Fig. 5s–u and Supplementary Table 4h). Therefore, activated fibroblasts have opposing intrinsic and extrinsic effects on reprogramming efficiency, and the relative proportions of activated and non-activated fibroblasts in cultures from old mice could underlie the variability in reprogramming efficiency.

To analyse whether extrinsic factors drive the variability in reprogramming efficiency between individual old cultures, we examined the difference in reprogramming efficiency between good and bad old fibroblast cultures, treated with their own conditioned medium or conditioned medium that was swapped between cultures (Fig. 3e and Extended Data Fig. 6a–c). Reprogramming pairs of good and bad old cultures with swapped conditioned medium reduced the difference between their reprogramming efficiencies (Fig. 3e) by more than 60% (Extended Data Fig. 6c). Extrinsic factors thus have a substantial role in the variability in reprogramming efficiency between old cultures, and intrinsic factors are likely to underlie the remainder of the effect.

We next tested whether cytokines contribute to the role of extrinsic factors on the variability between mice. IL-6, TNF and IL-1β, which are all secreted by old fibroblast cultures, affected reprogramming in opposing directions: IL-6 enhanced reprogramming efficiency (as previously reported[21]), whereas TNF and IL-1β impaired reprogramming efficiency in young and old fibroblasts (Extended Data Fig. 6d–i). Consistently, blocking IL-6 with an antibody reduced reprogramming efficiency, whereas blocking TNF improved it (Extended Data Fig. 6j, k). To determine whether IL-6 and TNF contributed to the variability between mice in reprogramming efficiency, we reprogrammed pairs of good old and bad old fibroblast cultures in their own conditioned medium, which was pretreated with IL-6- or TNF-blocking antibodies. While blocking IL-6 had a minor effect, blocking TNF reduced the difference in reprogramming efficiency between pairs of good old and bad old cultures (Fig. 3f) by more than 40% (Extended Data Fig. 6l–n). The IL-6:TNF ratio correlated with reprogramming efficiency (Fig. 3g and Extended Data Fig. 6o–q). Hence, the proportions of activated and non-activated fibroblasts, and the ratio of inflammatory cytokines that they secrete (for example, IL-6 and TNF), could drive the variability between fibroblast cultures of different old mice (Fig. 3h).

Fibroblasts are critical for wound healing in vivo[2,3,14,15]. Although the influence of ageing on wound healing has been examined[2,15,22,23], the variability of this response is not known. We assessed the rate of healing in wounds on the ears of young and old mice (Fig. 4a). While the median wound healing rate was not significantly affected by age, there was an increased variability in wound healing rate between old mice, with some old mice healing faster and some slower than young mice (Fig. 4a and Extended Data Fig. 7a–g).

To determine the overall fibroblast composition in wounds from young and old mice, we performed single-cell RNA-seq on FACS-sorted fibroblasts pooled from the wounds of 10 young or 10 old mice, 7 days after the induction of wounds—irrespective of wound healing rates (Fig. 4b and Extended Data Fig. 7c, d). Fibroblast composition changed in wounds from old mice in vivo (Fig. 4b), with subpopulations of fibroblasts exhibiting signatures of fibroblast activation and increased cytokine signalling (Extended Data Fig. 8a–f).

We next performed single-cell RNA-seq on all cells from the wounds of old mice with slow- or fast-healing trajectories (Fig. 4c and Extended Data Fig. 8g–i). Although epithelial cells were not identified (perhaps owing to the isolation protocol or wound composition and as previously reported[14]), fibroblasts, endothelial cells and immune cells were identified (Fig. 4c and Extended Data Fig. 8j). Notably, fibroblasts were more abundant in wounds of slow-healing old mice, whereas immune cells were more abundant in wounds of fast-healing old mice (Fig. 4c and Supplementary Table 5e). Although the number of mice is low and differences in the composition of cells could also be influenced by wound stage and isolation properties, fibroblast populations may therefore be associated with distinct wound healing trajectories.

Clustering using both Seurat and pathway and gene set overdispersion analysis (PAGODA) on wound fibroblasts from slow-healing or fast-healing old mice identified three main subpopulations (A, B and C) that were enriched in different aspects of fibroblast activation (Fig. 4d and Extended Data Fig. 9d, e; for a combined analysis of both single-cell RNA-seq datasets, see Extended Data Fig. 9h–l). Whereas fibroblast subpopulation A was present in wounds of both slow- and fast-healing mice, fibroblast subpopulation B was more abundant in wounds of fast-healing old mice and exhibited increased cytokine expression and signalling (for example, *Tnf*; Fig. 4d, Extended Data Fig. 9d, f, k and Supplementary Table 5f, g). Thus, TNF is associated with fast wound healing in vivo and bad reprogramming in vitro (fast wound healing might lead to fibrosis, which is detrimental). By contrast, fibroblast subpopulation C was more abundant in wounds from slow-healing old mice and exhibited higher expression of other cytokines (for example, *Ccl11*) and the transcription factor *Ebf2* (Fig. 4d, Extended Data Fig. 9d–g, k and Supplementary Table 5f, g). Activated fibroblast subpopulations with distinct cytokine profiles (for example, TNF compared to IL-6 or CCL11) may therefore be associated with increased variability in reprogramming in vitro and wound healing trajectories in old mice.

Our study shows that ageing is associated with an increased variability between mice in cellular reprogramming in vitro and in wound healing in vivo, perhaps reflecting different ageing trajectories. Increased variability is emerging as common feature of ageing[5–8], and we identify inflammatory cytokines, including TNF, as key contributing factors to variability in reprogramming efficiency (although other intrinsic and extrinsic factors may also exist). Cytokine signalling may also regulate the variability in other ageing phenotypes, including wound healing. Dermal fibroblasts have been shown to lose cellular identity and acquire adipogenic traits during ageing[19], and this increased cellular heterogeneity could also contribute to the differences between individual mice. As fibroblasts exhibit tissue-specific properties, variability in distinct tissues may differentially increase with age.

A subpopulation of activated fibroblasts could be a source of chronic inflammation in old individuals and contribute to immune cell recruitment[3,14,15,20]. Activated fibroblasts (which proliferate) and senescent fibroblasts (which show permanent cell cycle arrest) secrete overlapping yet distinct sets of cytokines[24] and may interact in a complex manner to influence reprogramming and wound healing. Wound healing is a major issue for elderly individuals, with either deficient wound healing (which can lead to ulcers) or excessive wound healing (which can lead to fibrosis)[2,3,15]. Changes in fibroblast subpopulations and cytokines with age could contribute to these pathologies and constitute targets for personalized strategies to restore functional wound healing in elderly individuals.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1658-5.


1. Franceschi, C. & Campisi, J. Chronic inflammation (inflammaging) and its potential contribution to age-associated diseases. *J. Gerontol. A Biol. Sci. Med. Sci.* **69**, S4–S9 (2014).
2. Eming, S. A., Martin, P. & Tomic-Canic, M. Wound repair and regeneration: mechanisms, signaling, and translation. *Sci. Transl. Med.* **6**, 265sr6 (2014).
3. Lynch, M. D. & Watt, F. M. Fibroblast heterogeneity: implications for human disease. *J. Clin. Invest.* **128**, 26–35 (2018).
4. Ocampo, A., Reddy, P. & Belmonte, J. C. I. Anti-aging strategies based on cellular reprogramming. *Trends Mol. Med.* **22**, 725–738 (2016).
5. Chaleckis, R., Murakami, I., Takada, J., Kondoh, H. & Yanagida, M. Individual variability in human blood metabolites identifies age-related differences. *Proc. Natl Acad. Sci. USA* **113**, 4252–4259 (2016).
6. Ong, M. L. & Holbrook, J. D. Novel region discovery method for Infinium 450K DNA methylation data reveals changes associated with aging in muscle and neuronal pathways. *Aging Cell* **13**, 142–155 (2014).
7. Li, R. et al. Linking inter-individual variability in functional brain connectivity to cognitive ability in elderly individuals. *Front. Aging Neurosci.* **9**, 385 (2017).
8. Cheung, P. et al. Single-cell chromatin modification profiling reveals increased epigenetic variations with aging. *Cell* **173**, 1385–1397 (2018).
9. Lapasset, L. et al. Rejuvenating senescent and centenarian human cells by reprogramming through the pluripotent state. *Genes Dev.* **25**, 2248–2253 (2011).

10. Li, H. et al. The *Ink4/Arf* locus is a barrier for iPS cell reprogramming. *Nature* **460**, 1136–1139 (2009).
11. Ravaioli, F., Bacalini, M. G., Franceschi, C. & Garagnani, P. Age-related epigenetic derangement upon reprogramming and differentiation of cells from the elderly. *Genes* **9**, 39 (2018).
12. Mosteiro, L. et al. Tissue damage and senescence provide critical signals for cellular reprogramming in vivo. *Science* **354**, aaf4445 (2016).
13. Somers, A. et al. Generation of transgene-free lung disease-specific human induced pluripotent stem cells using a single excisable lentiviral stem cell cassette. *Stem Cells* **28**, 1728–1740 (2010).
14. Guerrero-Juarez, C. F. et al. Single-cell analysis reveals fibroblast heterogeneity and myeloid-derived adipocyte progenitors in murine skin wounds. *Nat. Commun.* **10**, 650 (2019).
15. Shook, B. A. et al. Myofibroblast proliferation and heterogeneity are supported by macrophages during skin repair. *Science* **362**, eaar2971 (2018).
16. Frenk, S. & Houseley, J. Gene expression hallmarks of cellular ageing. *Biogerontology* **19**, 547–566 (2018).
17. Bahar, R. et al. Increased cell-to-cell variation in gene expression in ageing mouse heart. *Nature* **441**, 1011–1014 (2006).
18. Wiley, C. D. et al. Analysis of individual cells identifies cell-to-cell variability following induction of cellular senescence. *Aging Cell* **16**, 1043–1050 (2017).
19. Salzer, M. C. et al. Identity noise and adipogenic traits characterize dermal fibroblast aging. *Cell* **175**, 1575–1590 (2018).
20. Croft, A. P. et al. Distinct fibroblast subsets drive inflammation and damage in arthritis. *Nature* **570**, 246–251 (2019).
21. Brady, J. J. et al. Early role for IL-6 signalling during generation of induced pluripotent stem cells revealed by heterokaryon RNA-seq. *Nat. Cell Biol.* **15**, 1244–1252 (2013).
22. Keyes, B. E. et al. Impaired epidermal to dendritic T cell signaling slows wound repair in aged skin. *Cell* **167**, 1323–1338 (2016).
23. Nishiguchi, M. A., Spencer, C. A., Leung, D. H. & Leung, T. H. Aging suppresses skin-derived circulating SDF1 to promote full-thickness tissue regeneration. *Cell Rep.* **24**, 3383–3392 (2018).
24. Coppé, J. P. et al. Senescence-associated secretory phenotypes reveal cell-nonautonomous functions of oncogenic RAS and the p53 tumor suppressor. *PLoS Biol.* **6**, e301 (2008).

# Methods

### Mice

All mice used in this study were male C57BL/6 mice. Mice of different ages (3–29 months) were obtained from the National Institute on Ageing (NIA) colony, and were acclimatized to the animal facility at Stanford University for at least 1 week before being processed. No live animals were censored. For most animal experiments, young and old mice were processed in an alternate manner rather than in two large groups, to minimize group effects, and no blinding was performed. All experimental procedures were approved by Stanford's Administrative Panel on Laboratory Animal Care and were in accordance with institutional and national guidelines. At Stanford University, all mice were housed in the Comparative Medicine Pavilion, and their care was monitored by the Veterinary Service Center at Stanford University under IACUC protocol 8661.

### Collection of blood and plasma from young and old mice

To assess the systemic changes associated with age, whole blood was collected from young and old mice by cardiac puncture into a tube containing EDTA (Thermo Fisher Scientific, AM9262) (for a final concentration of 5 mM EDTA per blood sample). Blood cell composition, including white and red blood cell, granulocyte, monocyte, lymphocyte and platelet counts were analysed with a Hemavet Multispecies Hematology Analyzer (CDC Technologies) according to the manufacturer's instructions. Plasma was prepared from whole blood samples by two consecutive centrifugation steps at 500 r.c.f. and 13,000 r.c.f., respectively, each for 10 min at room temperature, and then aliquoted and stored at −80 °C for cytokine profiling (see 'Cytokine profiling analysis on plasma and conditioned medium using Luminex multi-analyte').

### Generation of primary cultures of fibroblasts from young and old mice

To investigate the effect of ageing on tissue fibroblasts, primary fibroblast cultures were established from the ears and lungs of young and old mice. To this end, the ears and lungs were cut into small fragments (approximately 1 $mm^2$) and digested in Dulbecco's modified Eagle medium (DMEM, Invitrogen, 11965-092) supplemented with 0.14 Wunsch units $ml^{-1}$ of Liberase TM (Roche, 05401127001) for 30–90 min. The fragments were washed with DMEM supplemented with 15% fetal bovine serum (FBS, Gibco, 16000-044, lots 551495 and 1551824) and plated on tissue culture plates with DMEM supplemented with 15% FBS and 1% penicillin–streptomycin–glutamine (PSQ) (Gibco, 10378). To isolate primary adult fibroblasts from the chest area, the skin on the chest was dissected from the animals, the subcutaneous fat and fascia were removed, and the tissues were incubated overnight at 4 °C with the epidermal layer of the skin facing down on top of a solution of 0.25% trypsin (Gibco, 25200-056). The following day, the epidermis was removed, tissues were cut into small fragments (approximately 1 $mm^2$) and treated with 1,000 U $ml^{-1}$ collagenase I (Gibco, 17100017) in DMEM for 60–90 min at 37 °C. Digested fragments were funnelled through a 70-μm nylon mesh (Fisher Scientific, 08-771-2), washed with fibroblast growth medium (DMEM supplemented with 10% FBS and 1% PSQ) and plated using the same medium. The cells were passaged once, before being aliquoted, frozen and stored in liquid nitrogen (passage 1.5). For all experiments, unless stated, fibroblasts were thawed (passage 2) and cultured at 37 °C in 5% $CO_2$ and 95% humidity in fibroblast growth medium. All experiments, unless specifically noted, were performed at passage 3.

### FACS analysis of primary fibroblasts

To determine the purity of the primary fibroblasts from young and old mice, FACS analysis was performed on fibroblast cultures at passage 3. FACS analysis was performed using an LSR II flow cytometer (BD Biosciences) and analysed using FlowJo v.10.0.7. For FACS analysis, fibroblasts were stained with phycoerythrin-conjugated CD140a (BioLegend, 135905), in combination with the following allophycocyanin-conjugated antibodies: B220 (eBioscience, 47-0452-82), CD3 (BD Pharmingen, 557597), Gr-1 (eBioscience, 17-5931-82), F4/80 (eBioscience, 17-4801-82), Siglec H (BioLegend, 129611), CD11c (eBioscience, 17-0114-82) and propidium iodide staining solution (BD Pharmingen).

### Generation of primary cultures of fibroblasts from young and old human individuals

To determine whether primary fibroblasts from humans also exhibit an inflammatory profile, we collected biopsies from humans at different ages. Stanford Human Subjects approval and informed consent was obtained before all study procedures (under protocol ID 25269, IRB 350). Biopsies were collected from male participants of different ages with four biological grandparents of Ashkenazi Jewish descent, generally healthy without thyroid disease, diabetes, immunodeficiency, ongoing cancer or autoimmune disease, and no history of poor wound healing (Supplementary Table 1g). A 4-mm punch biopsy of pre-auricular skin was obtained after injection of 1% lidocaine with epinephrine (1:1,000,000). Skin biopsies were rinsed with PBS, cut into smaller fragments (around 1 $mm^2$) and plated into a dry 6-well tissue-culture plate. Excess PBS was removed, and fibroblast growth medium (DMEM supplemented with 10% FBS and 1% PSQ) was added. Tissues were incubated at 37 °C in 5% $CO_2$ and 95% humidity. After 24 h, tissues were supplemented with fibroblast growth medium, and the medium was changed every 3–4 days. The cells were passaged once, before being aliquoted, frozen and stored in liquid nitrogen (passage 1.5).

### Cytokine profiling analysis on plasma and conditioned medium using Luminex multi-analyte

We examined the effect of ageing on the inflammatory profiles by performing cytokine profiling on plasma and conditioned medium from fibroblast and iPS cell cultures from young and old mice. Plasma was collected as described above (Supplementary Table 1a). Conditioned medium from young and old mouse (ear and lung) and human (skin) fibroblasts was collected 48 h after plating from 150,000–200,000 primary fibroblasts (passage 3 or 33) plated in a 6-cm dish with 2 ml of fibroblast growth medium (Supplementary Table 1b–d). Conditioned medium from iPS cells (passage 23; Extended Data Fig. 1g and Supplementary Table 1e) was collected 24 h after plating from 500,000 cells maintained in serum- and feeder-free culture conditions in 2i medium (see 'Cytokine profiling analysis on plasma and conditioned medium using Luminex multi-analyte' for more information). Conditioned medium from cultures of $THY1^+PDGFR\alpha^+$ and $THY1^-PDGFR\alpha^+$ FACS-sorted young and old fibroblasts (passages 4–6, see 'FACS and analysis of primary fibroblasts' for FACS sorting protocol) was collected 24 h after plating from 0.5–1 million cells plated in a 15-cm dish with 20 ml of medium (Supplementary Table 4h). Conditioned medium was collected, centrifuged at 10,000 r.c.f. for 10 min at room temperature, aliquoted and stored at −80 °C. For all of these conditions, cell numbers were determined for each plate by counting on haemocytometer for normalization purposes. In addition, cell-free medium was used to assess background fluorescence. All cytokine profiling was performed by the Stanford Human Immune Monitoring Center using a Luminex mouse 38-plex or a human 62-plex analyte platform (eBiosciences/Affymetrix) that detects 38 or 62 secreted proteins, respectively.

All plasma samples were measured in technical duplicates and all conditioned medium samples were measured in single technical replicates as per recommendation of the Human Immune Monitoring Center at Stanford University. All of our analyses were performed using mean fluorescence intensity (MFI) values, because converting MFI to clinically relevant measures (such as pg $ml^{-1}$) can introduce a degree of error[60]. We report pg $ml^{-1}$ conversions in Supplementary Table 1a–e to facilitate comparison with existing literature. To compare values across plates and independent experiments, the MFI values were normalized to the median of young (3 months) within each experiment, generating fold

change values. In addition, the conditioned medium levels were normalized to the cell number of the same dish. Two plasma samples from old mice were discarded, as the coefficient of variation was >20% for most of the cytokines measured between the two technical replica for these two plasma samples. Ranked fold changes in cytokine levels were calculated by multiplying the $\log_2$-transformed fold median change (old/young) with the $-\log_{10}(P)$ values. Similarly, ranked Spearman $\rho$ correlations were calculated by multiplying the Spearman $\rho$ values with $-\log_{10}(P)$ values.

### Lentiviral production for reprogramming

To induce reprogramming in fibroblasts and generate iPS cells, we used the lentiviral vector 4F STEMCAA–loxP, containing a floxed version of EF1α–STEMCCA enabling the expression of human *OCT4*, *KLF4*, *SOX2* and *MYC*. Lentiviruses were produced in human embryonic kidney 293T (HEK293T, ATCC, CRL-11268) packaging cells. The HEK293T cell line was not authenticated in-house, but mycoplasma testing was conducted at regular intervals (every 2–3 months). The day before transfection, $9 \times 10^6$ HEK293T cells were plated in a 10-cm dish in HEK293T medium (DMEM supplemented with 10% FBS, 1% PSQ). The next day, the cells were transfected as follows: 100 µl of 1 mg ml$^{-1}$ polyethylenimine (PEI; Polysciences, 23966-2, linear 25 kDa) was added to 2 ml of DMEM and incubated for 10 min at room temperature. The lentiviral vector of interest (20 µg) was mixed with lentiviral packaging vectors (1 µg of pHDM-tat1b (PlasmID), 1 µg of pRC-CMV-rev1b (PlasmID), 1 µg of pHDM-Hgpm2 (PlasmID)) and envelope vector (2 µg of HDM-VSV-G (PlasmID), added to the PEI–DMEM mixture and incubated for 15 min at room temperature. The PEI–DMEM–DNA mixture was then added dropwise to the HEK293T cells, and 12 h after transfection the medium was replaced with 8 ml fresh HEK293T medium. Viral supernatants were collected at 24 and 36 h after transfection, centrifuged at 3,000 r.p.m. for 15 min, and carefully transferred into a fresh tube, after which 0.7 ml of the crude virus supernatant was used to reprogram primary fibroblasts (see 'Reprogramming of young and old fibroblasts to iPS cells and characterization of the iPS cells').

### Reprogramming of young and old fibroblasts to iPS cells and characterization of the iPS cells

We generated iPS cell lines from three independent young fibroblast cultures and from three independent old fibroblast cultures (Supplementary Table 1f). Reprogramming of primary fibroblasts was induced as follows: 100,000 primary fibroblasts at passage 3 were plated in a well of a 6-well plate, and were infected 24 and 36 h after plating with 0.7 ml crude virus supernatant mixed with 8 µg ml$^{-1}$ polybrene (Sigma-Aldrich, H9268-5G). Next, 48 h after plating (12 h after the last round of infection), the infected primary fibroblasts were plated at a density of 5,000 cells on a 10-cm dish containing $1.5 \times 10^6$ γ-irradiated feeder cells (mouse embryonic fibroblasts (MEFs)). Cells were maintained in fibroblast growth medium for 7 days, and then switched to mouse embryonic stem (mES) cell medium, consisting of DMEM, GlutaMax (Life Technologies, 10569-010), 15% FBS, 1% PSQ, $5 \times 10^5$ units of leukaemia inhibitory factor (EMD Millipore, ESG10007), 1% MEM nonessential amino acids (Gibco, 11140-050) and 0.0008% β-mercaptoethanol (Sigma-Aldrich, M-7522). On days 13–15 after infection, colonies with a distinct mES cell morphology were manually picked from 10-cm dishes and each iPS cell clone was transferred into a well of a 96-well plate (primary plate) in the presence of γ-irradiated MEFs. A minimum of 24 iPS cell clones per parental fibroblast line were picked and replicates of each 96-well primary plate were created. These replicate plates were used to evaluate the number of viral integrations in each clone, whereas the primary plates were temporarily frozen and stored at −80 °C. To determine the number of viral integrations, on-plate genomic DNA extractions were performed as previously described[25], and the Mouse TaqMan Copy Number Reference Assays from Thermo Fisher was used to estimate the number of viral integrations from the genomic DNA extracted. A TaqMan probe targeting the human *KLF4* gene (FAM dye labelled) was used because the 4F STEMCAA–loxP vector contains the human version of the reprogramming factors (Life Technologies, 4331182). A TaqMan probe targeting the mouse transferrin receptor gene (*Tfrc*), which is known to be encoded by a single gene in the mouse genome, was used as the reference (VIC dye labelled) (Life Technologies, 4458366). Only iPS cell clones with an estimated viral integration number equal to or lower than 3 were chosen for further analysis.

For 13 of these lines, we generated transgene-free iPS cell lines by excising the reprogramming factor construct and performed long-term passaging (until passage 23), as this is known to improve the pluripotency state[26]. To this end, primary plates were quickly thawed and the iPS cell clones were transferred into a fresh 96-well plate in the presence of γ-irradiated MEFs, and subsequently expanded. At passage 10, the integrated 4F STEMCCA lentiviral construct was excised using Cre-recombinase expressed under the CAG promoter (pCAG-Cre)[13]. The pCAG-Cre construct was transfected using a Mouse ES Cell Nucleofector Kit (LONZA, V4XP-3012) according to the manufacturer's instructions. Transfected cells were then resuspended in mES cell medium, plated on feeder cells at a very low density in a 10-cm dish (500 cells per dish) and cultured in mES cell medium until colonies appeared. For each iPS cell clone, multiple subclones were isolated and expanded. The efficiency of Cre-recombinase excision was assessed by PCR using the Mouse TaqMan Copy Number Reference Assays as described above. Only transgene-free iPS cell clones were further characterized. iPS cell lines were maintained on ES cell medium for 10 passages after excision, before being adapted to serum- and feeder-free culture conditions in 2i medium according to the CReM Boston University ES cell culture protocols (http://www.bu.edu/dbin/stemcells/protocols.php). All molecular characterizations of the iPS cell lines were performed at passage 23, including the inflammatory, transcriptomic and metabolomics profiling (Extended Data Fig. 3).

To assess whether the derived iPS cell lines could give rise to cell types from all three germ layers after formation of embryoid bodies, we induced the formation of embryoid bodies. In brief, iPS cells at passage 23 were incubated with accutase (EMD Millipore) for 5 min at 37 °C to obtain a single-cell suspension and 10 ml of the iPS cell suspension at a density of $10^3$ cells per ml was plated on ultralow attachment plates (Corning). Cells were allowed to form embryoid bodies. After 4 days, embryoid bodies were transferred into regular tissue-culture-grade plates in DMEM high glucose supplemented with 10% FBS, 100 U ml$^{-1}$ penicillin and 100 µg ml$^{-1}$ streptomycin (Gibco), and embryoid bodies were allowed to differentiate. At day 14 after embryoid body differentiation, differentiated cells were collected and analysed by qRT–PCR for the expression of endodermal, mesodermal and ectodermal markers (primer sequences are listed in Supplementary Table 6a).

### RT–qPCR on iPS cells and differentiated cells from embryoid bodies

To assess the expression of specific genes in iPS cells and in differentiated cells from embryoid bodies, RNA purification and cDNA synthesis was performed. To this end, total RNA was isolated using the RNeasy RNA Purification Kit (QIAGEN) and 0.5–1 µg of RNA was reverse-transcribed using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems) according to the manufacturer's instructions. cDNA was used for RT–qPCR on the BioRad iCycler using iQ SYBR Green Mix (BioRad). *Hprt1* was used as housekeeping gene for normalization. All primer sequences are listed in Supplementary Table 6a.

### Assessment of reprogramming efficiency

To determine the impact of ageing on iPS cell generation, reprogramming efficiency was quantified using a 96-well assay as previously described[27]. In brief, reprogramming was induced as described above. Then, 48 h after plating (12 h after the last round of lentiviral infection), the infected primary fibroblasts were plated at a density of 20–40 cells per well into 96-well plates containing 1,000 γ-irradiated feeder MEFs per well. In experiments using cytokines and conditioned medium, 0.1%

gelatin-coated plates (Tribec Science, TBS8004) without feeder cells were used to avoid confounding factors from the feeder cells.

Infected primary fibroblasts were maintained on fibroblast growth medium until day 7 after plating and then switched to mES cell medium until day 13–15. For experiments assessing the effect of conditioned medium on reprogramming efficiency, fresh conditioned medium was collected from 10–25-cm dishes in which cells were grown in parallel, centrifuged 10,000 r.c.f. for 10 min at room temperature and added every day, starting from day 1 of replating into 96-well plates until the end of experiment. For experiments testing the influence of specific cytokines on reprogramming efficiency, fresh medium with the indicated cytokine was added every day until the switch to iPS cell medium.

To assess reprogramming efficiency, staining with AP (an early marker of pluripotency[10]) was performed by fixing the cells in 4% paraformaldehyde (Santa Cruz Biotechnology, sc-281692) for 15 min at room temperature, washing with citrate solution (Sigma-Aldrich, 3861) and subsequent staining with prepared diazonium salt solution (Sigma-Aldrich, 851) with napthol (Sigma-Aldrich, 855) overnight. Quantification was performed by counting the number of wells containing at least one $AP^+$ colony. To complement AP staining, we also used staining with stage-specific embryonic antigen 1 (SSEA1) (a later marker of pluripotency[10]). SSEA1 staining was performed using StainAlive mouse anti-mouse antibody (Stemgent, 09-0067) according to the manufacturer's recommendations. Quantification was performed by counting the number of wells containing at least one $SSEA1^+$ colony using a Zeiss inverted microscope (Zeiss AxioVision A10).

Reprogramming efficiency was calculated as the number of $AP^+$ or $SSEA1^+$ clones, divided by the number of cells plated, and multiplied by the efficiency of viral infection (see 'Immunofluorescence staining of reprogramming factors and pluripotency markers'). To compare reprogramming efficiencies across plates (and independent experiments), the reprogramming efficiencies of all individual cultures were normalized to the median reprogramming efficiency of young cultures within a given experiment. Statistical differences in variance in reprogramming efficiency between the age groups were calculated using the non-parametric Fligner–Killeen test using R v.3.3.0. To assess whether the increased variability in reprogramming efficiency with age was introduced by pooling multiple cohorts, we performed a permutation test in which the null distribution was estimated by randomly assigning the age groups to the observed reprogramming efficiencies for individual cultures within each cohort, and the mean difference in standard deviation between young and old cells was calculated across the cohorts. This was repeated 1,000 times, and the *P* value was calculated as the percentage of differences greater than or equal to the actual observed difference in standard deviation. This approach indicated that the increased variability in reprogramming efficiency with age is not simply caused by the pooling of multiple cohorts ($P < 0.001$).

### Immunofluorescence staining of reprogramming factors and pluripotency markers

For immunofluorescence staining of pluripotency markers, cells were fixed in 4% paraformaldehyde for 15 min at room temperature, then permeabilized with 0.5% Triton X-100 for 10 min, blocked in blocking solution (2% bovine serum albumin (BSA), 5% glycerol, 0.2% Tween-20, 0.1% sodium azide in PBS) for 1 h, followed by incubation with primary antibodies. The following antibodies were used for immunofluorescence: rabbit anti-OCT3/4 (Santa Cruz Biotechnology, sc9081), SSEA1 StainAlive mouse anti-mouse antibody (Stemgent, 09-0067) and rabbit anti-SOX2 (Santa Cruz Biotechnology, sc17320). The nuclei were stained with DAPI (Life Technologies). Cells were imaged using a Zeiss inverted microscope (Zeiss AxioVision A10) with AxioVision v.4.7.2 software. For calculations of the infection efficiency, 5–10 images were randomly taken per sample and uploaded in ImageJ (v.1.46r), and the infection efficiency was calculated by dividing the number of $OCT4^+$ cells by the total number of cells (as determined by DAPI staining).

### Reprogramming of young and old fibroblasts to induced neurons

To determine the ability of young and old primary fibroblasts to reprogram to induced neurons, induced neuron reprogramming was induced as previously described[28]. In brief, young and old fibroblast cultures (passage 3) were plated at a density of 60,000 cells per well in a 12-well plate. The following day, the fibroblasts were infected as described above with lentiviruses carrying TetO-FUW-ASCL1 (Addgene, 27150), TetO-FUW-BRN2 (Addgene, 27151), TetO-FUW-MYT1L (Addgene, 27152) and FUW-rtTA (Addgene, 20342). The next day, doxycycline (2 μg $ml^{-1}$, Sigma-Aldrich) in fibroblast growth medium was added to the wells. Medium was changed to neuronal medium (N2, B27, DMEM/F12 (Invitrogen), 1.6 ml insulin (6.25 mg $ml^{-1}$, Sigma-Aldrich)) and doxycycline (2 μg $ml^{-1}$) two days after the first doxycycline induction. Subsequently, neuronal medium was changed every three days. To determine the number of induced neurons at day 7 for each fibroblast culture, the cells were digested using 0.25% trypsin (Invitrogen) at 37 °C for 5 min, and all cells were subjected to magnetic activated cell sorting (MACS) to select for APC-conjugated PSA-$NCAM^+$ cells (Miltenyi, 130-093-273), according to the manufacturer's instructions. The number of PSA-$NCAM^+$ cells for each fibroblast culture was counted manually using a haemocytometer. The reprogramming efficiency for each line was obtained by dividing the total number of PSA-$NCAM^+$ cells obtained at day 7 by the number of fibroblasts plated. The ability of the primary fibroblast cultures to undergo induced neuron and iPS cell reprogramming was assessed in parallel. Note that in this comparison, infection efficiency was not assessed and hence not included in the calculation of reprogramming efficiency.

### RNA-seq analysis

To profile transcriptomic changes in primary fibroblast cultures with age and after iPS cell reprogramming, total RNA was isolated from passage 3 fibroblasts and passage 23 iPS cells using the RNeasy kit (QIAGEN) according to the manufacturer's instructions. Total RNA (150 ng) was used to prepare RNA-seq libraries using the Encore Complete RNA-seq library kit (Nugen Technology, 0333), according to the manufacturer's instructions. Libraries were sequenced on HiSeq 2000 (2 × 10 bp paired-end reads, Illumina).

Quality and adaptor trimming of the Fastq files was performed using TrimGalore v.0.2.8, retaining reads with a minimum Phred score of 15. The trimmed reads were mapped to the mouse genome (mm9 build) using TopHat (v.2.0.8b). Reads per genes were counted using HTSeq (v.0.6.1). As annotation file, we used the genes.gtf downloaded from UCSC on 6 March 2013. Gene expression was analysed using DESeq2 (v.1.20.0). For differential expression analysis of fibroblasts, batch effect was accounted for by including a batch variable into the DESeq2 model (see Supplementary Table 2a). Genes with >0.3 fragments per kb of transcript per million mapped reads in at least one sample within a particular analysis, were considered expressed and included in the analysis. Heat maps, hierarchical clustering and principal component analysis (PCA) were performed on VST-transformed values (implemented in DESeq2). Genes were considered significantly differentially expressed if they had FDR-adjusted values of $P < 0.05$ and an absolute fold change >1.5, unless stated otherwise. Publicly available datasets were downloaded from the GEO database (Supplementary Table 6b) and processed as described above. Note these following RNA-seq samples were excluded from further analyses: (1) two old and three middle-aged RNA-seq libraries as they lacked any young samples, and hence batch-effects could not be corrected for; (2) RNA-seq libraries from one good old and one bad old fibroblast cultures as their reprogramming efficiency could not be confirmed across several independent experiments; (3) RNA-seq libraries from 2 iPS cell lines (out of 13 total) failed at the quality-control stage because they showed large differences (for example, in number of reads mapped) from the rest of the samples (Supplementary Table 1f).

Pathway enrichment analysis was performed using one-sided Fisher's exact tests, testing for the overrepresentation of significantly differentially expressed genes in a given gene list. As background, all of the genes that were considered expressed (see above) were used. *P* values adjusted for multiple hypothesis testing using Benjamini–Hochberg correction, and FDR-adjusted $P = 0.05$ was set as upper threshold. In Extended Data Fig. 2m, analysis of gene set enrichment was conducted using the gene set enrichment analysis (v.2.2.2) tool. For this analysis, the VST-transformed values (derived from DESeq2) were used, and enrichment statistics were calculated using the 'classic' method parameter. Nominal *P* values were calculated based on 10,000 permutations. In Fig. 2e and Extended Data Figs. 2w, 5j, 7e, analysis of gene set enrichment was conducted by calculating the arithmetic mean of gene-wise test statistics (Wald test statistic from differential expression analysis using DESeq2) per gene set. To calculate a *P* value for each gene set, we constructed a null distribution of test statistics by sampling 10,000 times *n* genes (*n* indicates the number of genes in the respective gene set) and calculating the mean of the test statistics for these genes. A gene set-wise *P* value was then calculated as the percentage of absolute (sampled) mean test statistics that were equal or greater than the absolute (observed) mean test statistic for that pathway. *P* values were corrected for multiple hypothesis testing using the Benjamini–Hochberg algorithm using FDR-adjusted $P = 0.05$ as threshold. This method for gene set enrichment analysis has been shown to outperform many commonly used methods[29]. KEGG, Gene Ontology (GO) terms were acquired from http://amp.pharm.mssm.edu/Enrichr/#stats.

Upstream regulator analysis was performed using ingenuity pathway analysis (IPA; QIAGEN) software, using the genes that passed the filter in our dataset as reference genome.

Motif analysis of promoter regions (−1,000 to +50 bp relative to the transcription start sites) of differentially expressed genes was performed using the Homer software (v.4.8)[30], using the genes that passed the filter in our dataset as background.

### Chromatin immunoprecipitation followed by sequencing and analysis of the epigenomic landscape

To profile changes in the epigenomic landscape of primary fibroblasts with age, we performed ChIP experiments using anti-H3K4me3 (Active Motif, 39159) and anti-H3K27me3 (Active Motif, 39536) antibodies. In brief, $1–2 \times 10^6$ fibroblasts were crosslinked with 1% formaldehyde for 10 min at room temperature, and formaldehyde was quenched by addition of glycine to a final concentration of 0.125 M. Chromatin was sonicated to an average size of 0.5–2 kb, using Bioruptor (Diagenode). A total of 5 μg of antibody was added to the sonicated chromatin and incubated overnight at 4 °C on a rotating platform. Subsequently, 10% of chromatin used for each ChIP reaction was retained as input DNA. Then, 100 μl of protein G Dynal magnetic beads were added to the ChIP reactions and incubated for an additional 4 h at 4 °C. Magnetic beads were washed, followed by reversal of crosslinks and DNA purification. Resultant ChIP DNA was dissolved in water. ChIP and input libraries were generated according to the Illumina protocol and sequenced as single-end 50 bp reads using the Illumina HiSeq 2000 platform.

For analysis, Fastq reads were quality-trimmed using the trim-galore software (v.0.2.1), with a Phred score threshold of 15 and a minimum remaining read length of 36 bp. Trimmed reads were mapped to the mm9 genome assembly using Bowtie v.0.12.7[31]. Duplicate reads were eliminated using the FIXSEQ software with default parameters[32]. ChIP–seq peaks were called in all samples using the MACS (v.2.08) software with default settings and the --broad option[33,34]. Input datasets were used as baseline.

To identify H3K4me3 and H3K27me3 ChIP–seq peaks with differential intensity in young compared to old or good old compared to bad old samples, we used the DiffBind R package (v.1.12.3)[35]. Reads were quantified in each sample over 'meta-peaks', that is, peaks called using pooled reads from one specific mark (H3K4me3 reads and H3K27me3 reads) over all samples. Meta-peaks help to best determine peak boundaries[36]. ChIP–seq read counts normalized to input reads counts by DiffBind were then analysed using the DESeq2 package (v.1.6.3)[37] to identify peaks with significantly different intensity. Hierarchical clustering and PCA were performed on VST-transformed values (implemented in DESeq2). The differential peak intensity and pathway analyses were restricted to the peaks that extended at least 100 bp into the promoter regions of the nearest genes (defined as transcriptional start site ±2,000 bp), and were performed as mentioned above (see 'RNA-seq analysis').

Broad H3K4me3 domains are genomic regions coated with H3K4me3 and are enriched at genes involved in cell identity and/or function[38]. To compare H3K4me3 breadth of samples across young and old samples, we used the approach described previously[38]. In brief, we used the H3K4me3 meta-peaks to compare the signal-to-noise ratio across samples. This revealed that sample number 2 for H3K4me3 from 3-month-old fibroblasts was the noisiest sample of the 5 samples. We therefore downsampled all other samples to match the coverage histogram of that specific sample. We then called peaks as described above using the calibrated files in MACS (v.2.08) and isolated the top 5% broadest H3K4me3 domains (broad H3K4me3 domains) from each peak file. We identified reproducible broad H3K4me3 domains by retaining only those that were present in all young or all old samples, and we restricted the analysis to those.

Bivalent domains are genomic regions coated with both H3K4me3 and H3K27me3[39,40]. To identify differential bivalent regions between young and old samples, the H3K4me3 and H3K27me3 peaks that are consistently present in young samples were compared to the ones that are consistently present in old samples. To define robust bivalently marked regions in each age group, we called H3K4me3 and H3K27me3 meta-peaks separately at each age. Then, at each age and for each mark, we identified peaks that were supported by all of the individual experimental samples (that is, reproducible peaks). Bivalent peaks were obtained by the intersection of H3K4me3 and H3K27me3 reproducible peaks in all young or old samples. Note that the pathway enrichment analysis was restricted to the bivalent domains in young that lose H3K27me3 in old and to the H3K4me3 peaks in young that gain H3K27me3 in old, as these domains are likely to exhibit altered expression of their associated genes.

Together, the epigenomic profiling identifies age-dependent changes in the epigenomic landscape (for example, H3K4me3 intensity and breadth) and reveals enrichment of pathways involved in activated fibroblasts, such as cytokines, extracellular matrix components and contractility-related features (Fig. 2d, Extended Data Fig. 2n–q, x and Supplementary Table 2g–l), corroborating the transcriptomic findings.

### Metabolomics analysis

To profile changes in metabolomics features in cultured fibroblasts with age, frozen cell pellets were mixed with 80% methanol (mass-spectrometry-grade) in a ratio of 10 μl per mg cell pellet (a million cells weighs roughly 13 mg). The suspension was then processed by three rounds of 1 min vortex at maximum speed, chilled briefly on ice. The mixture was incubated at 4 °C for 20 min before centrifuging at 20,000*g* for 20 min at 4 °C. The supernatants were used as metabolite extracts for liquid chromatography–mass spectrometry analysis. For liquid chromatography–mass spectrometry analysis, the metabolite extracts were transferred to 150 μl deactivated glass insert housed in 2-ml brown mass-spectrometry vials (Waters). A chemical standard solution (for quality control) was prepared from a synthetic complete mixture from Sigma-Aldrich (Y1501) at a concentration of 19 μg $ml^{-1}$ in 80% mass-spectrometry-grade methanol (Fisher Scientific). Metabolite extracts were analysed in a platform that consisted of a Waters UPLC-coupled Exactive Orbitrap mass spectrometer (Thermo), using an OPD2 HP-4B column (4.6 mm × 50 mm) and an OPD2HP-4A guard column (Shodex). The column temperature was maintained at 45 °C. In brief, 5 μl of each sample maintained at 4 °C was loaded by the autosampler in partial loop mode 3 times in the positive mode and 3 times in the negative mode.

The binary mobile phase solvents were: A, 10 mM $NH_4OAc$ in 10:90 acetonitrile:water; B, 10 mM $NH_4OAc$ in 90:10 acetonitrile:water. Both solvents were modified with 10 mM HOAc for positive-mode acquisition or 10 mM $NH_4OH$ for negative mode. The 30-min gradient for both modes was set as: flow rate, 0.1 ml $min^{-1}$; 0–15 min, 99% A; 15–20.5 min, 99% to 1% A; 20.5–25 min, 1% A; 25–25.5 min, 1% to 99% A; 25.5–30 min, 99% A. The mass-spectrometry acquisition was in profile mode and performed with an electrospray ionization probe, operating with capillary temperature at 275 °C, sheath gas at 40 units, spray voltage at 3.5 kV for positive mode and 3.1 kV for negative mode, capillary voltage at 30 V, tube lens voltage at 120 V and skimmer voltage at 20 V. The mass scanning used 100,000 mass resolution, high dynamic range for AGC target, 500 ms as maximum inject time, and 75–1,200 $m/z$ as the scan range. The system was operated by Thermo Xcalibur v.2.1 software. The raw data files generated from liquid chromatography–mass spectrometry were centroided with PAVA program[41] and converted to mzXML format. Mass feature extraction was performed with XCMS v.1.30.3[42]. Differential analysis was performed on signal intensity values derived from XCMS using the nonparametric Wilcoxon rank-sum test for positive and negative mode separately and adjusted for multiple hypothesis testing using $q$ value correction using the R package $q$ value (v.2.0.0). The mass features that were found significantly different were manually searched against the Metlin metabolite database (29381867) using 5 ppm mass accuracy. Retention time matching with compounds in the standard mixture was also performed for a subset of the metabolite hits. Before PCA and hierarchical clustering analysis, signal intensity values derived from XCMS were range-scaled[43]. Pathway analysis was performed using the integrated pathway analysis tool in the MetaboAnalyst 3.0 software[44], using all putatively identified metabolites that were found significantly different (FDR-adjusted $P < 0.05$, absolute fold change >1.5) together with all differentially expressed genes from the transcriptomic analysis (see above).

Collectively, this metabolomics profiling uncovers changes in arginine and proline metabolism (Extended Data Fig. 2r–t and Supplementary Table 2m), which has been implicated in the regulation of inflammatory cytokines and extracellular matrix synthesis[45–47] (Extended Data Fig. 2t), consistent with the characteristics of activated fibroblasts.

### Single-cell RNA-seq analysis of primary cultures of fibroblasts

To assess the cell composition and heterogeneity of primary fibroblast cultures, single cells were isolated from three young and three old fibroblast cultures at passage 3 (see Supplementary Table 3g). In brief, 20 single cells per culture were isolated manually by picking isolated cells under a Zeiss inverted microscope (Zeiss AxioVision A10). Single-cell RNA-seq libraries were generated using SMARTer Ultra Low Input RNA Kit for Sequencing v.3 (Clontech, 634853), according to the manufacturer's instructions. Single cells were directly lysed in 2.5 µl of Clontech reaction buffer and the volume was brought up to 10 µl with sterile water. First-strand cDNA synthesis was carried out in 96-well PCR plates as follows: 1 µl of 3′ SMART CDS Primer II A (24 µM) was added and the resulting mix was incubated in a preheated thermocycler at 72 °C for 3 min and then held at 4 °C. Next, 7.5 µl of first-strand master mix was added (SMARTScribe Reverse Transcriptase, 5× First-Strand Buffer, dNTP Mix and SMARTer IIA Oligonucleotides), mixed and incubated at 42 °C for 90 min and 70 °C for 10 min. Finally, the first-strand cDNA was purified with SPRI Ampure XP beads; 36 µl of the SPRI beads was added to each 20-µl single-stranded cDNA sample, mixed and incubated for 8 min at room temperature. The samples were placed on a Promega MagnaBot II magnetic separation device, the supernatant was discarded, and the single-stranded cDNA sample bound to the beads was directly used for double-stranded cDNA generation. Next, 50 µl of PCR master mix was added to each sample and mixed. Plates were placed in a preheated thermal cycler with a heated lid using the following program: 95 °C for 1 min, 18 cycles of 95 °C for 30 s, 65 °C for 30 s, 68 °C for 6 min, followed by 72 °C for 10 min and hold on 4 °C. Amplified double-stranded cDNA was purified using SPRI Ampure Beads (Beckam Coulter), eluted in 12 µl of purification buffer and kept in −20 °C. The quantity and quality of 1 µl of the amplified purified double-stranded cDNA were measured using the Agilent 2100 BioAnalyzer and Agilent's High Sensitivity DNA Kit (Agilent, 5067-4626). Double-stranded cDNA libraries for which the BioAnalyzer results showed no contamination, a distinct peak at around 2,000 bp and with approximately 2–7 ng of cDNA were selected. This resulted in 8–12 single-cell cDNA samples from each culture. To generate RNA-seq libraries, we next used Nextera XT DNA Library Preparation kit and Nextera XT Index kit (Illumina, FC-131-1096 and FC-131-1002, respectively). In brief, 5 µl purified double-stranded cDNA (around 1 ng total) from the previous step was added into each sample well of a 96-well plate, and 10 µl Tagmentation (TD) buffer was added into each sample and mixed gently. Next, 5 µl amplicon tagmentation mix was added to the wells and mixed gently. The 96-well plate was sealed and placed in a thermal cycler and incubated at 55 °C for 5 min and held at 10 °C. The Tn5 transposase was inactivated by adding 5 µl of Neutralization buffer. The tagmented DNA was then amplified by adding 15 µl of Nextera PCR Master Mix, 5 µl index 1 primers (i7) and 5 µl index 2 primers (i5) to each sample. The final PCR was performed using the following program on a thermal cycler: 72 °C for 3 min, 95 °C for 30 s, 12 cycles of: 95 °C for 10 s, 55 °C for 30 s, 72 °C or 30 s and 72 °C for 10 min. The PCR products were then purified with Ampure beads. The final libraries were assessed using the Agilent 2100 BioAnalyzer and Agilent's High Sensitivity DNA Kit. We generated three pooled libraries and sequenced them on three lanes of Illumina HiSeq 2000 paired-end 2 × 101-bp sequencing reads. Quality and adaptor trimming of the Fastq files was performed using TrimGalore v.0.2.8, retaining reads with a minimum Phred score of 15. The trimmed reads were mapped to the mouse genome (mm9 build) using TopHat (v.2.0.8b). Reads per genes were counted using HTSeq (v.0.6.1). As annotation file, we used the genes.gtf downloaded from UCSC on 6 March 2013. On average, 7,000 genes were expressed per cell. Genes with at least 10 reads in 3 single cells were considered expressed. Heat maps, hierarchical clustering and PCA were performed on VST values (implemented in DESeq2).

### *t*-SNE and PAGODA analysis of single-cell RNA-seq data from cultured cells

To analyse the single-cell RNA-seq data, we performed *t*-SNE clustering using the Rtsne R package (v.0.14). Single-cell RNA-seq data were analysed using PAGODA[48]. PAGODA identifies pathways and sets of genes that are overdispersed in the data and separates the cells based on their expression patterns. We applied PAGODA to the raw counts of all genes that were considered to be expressed[48]. For gene sets, we used all KEGG pathways as well as an 'in vitro fibroblast ageing' gene set that we defined from comparing the population RNA-seq data from young and old fibroblast cultures (Supplementary Table 2b). In addition, we used the list of 'fibroblast activation' genes, which are genes that have previously been associated with fibroblast activation (Supplementary Table 2f). We used the PAGODA pipeline with default parameters, unless stated otherwise, and used the SCDE package v.0.99.1 in R v.3.2.2. PAGODA revealed a relatively strong cell clustering by KEGG cell cycle as well as two de novo gene sets (clusters 37 and 119; Extended Data Fig. 4b), consisting of many cell-cycle-related genes. We accounted for this cell cycle aspect of heterogeneity using the pagoda.subtract.aspect() method (see Supplementary Table 3h for the lists of genes in these gene sets). After accounting for cell cycle phases, PAGODA identified 74 KEGG pathways, 8 de novo gene sets and the in vitro fibroblast ageing and fibroblast activation signatures as significantly overdispersed in the dataset (Extended Data Fig. 4c).

### Immunofluorescence staining of fibroblast activation markers and EdU incorporation

Immunofluorescence staining was performed as described in 'Immunofluorescence staining of reprogramming factors and pluripotency

markers'. The following antibody was used for immunofluorescence: mouse anti-αSMA (Abcam, ab7817). The nuclei were stained with DAPI (Life Technologies). Cells were imaged using a Zeiss inverted microscope (Zeiss AxioVision A10) with AxioVision v.4.7.2 software.

EdU (5-ethynyl-2′-deoxyuridine) incorporation in fibroblast cultures was visualized using the Click-iT EdU Plus Alexa Fluor 594 Imaging Kit (Invitrogen, C10639). Fibroblasts were plated onto glass coverslips (Bellco Glass, 194310012A) in wells of 24-well plates at a density of 20,000 cells per well. After allowing the cells to attach overnight, fibroblasts were incubated in medium containing EdU (10 μM) for 4 h. Cells were then fixed (4% paraformaldehyde in PBS) and permeabilized (0.1% Triton X-100 in PBS). EdU was detected by click reaction according to the manufacturer's instructions. Cells were incubated in blocking buffer (2% BSA, 5% glycerol, 0.2% Tween-20, 0.1% sodium azide in MilliQ water) and stained with Alexa Fluor 488-conjugated anti-αSMA (Abcam, ab184675). Coverslips were mounted onto slides using ProLong Gold with DAPI (Invitrogen, P36931) and imaged on a Nikon Eclipse Ti/Andor CSU-W1 spinning disk confocal microscope using Andor Zyla and NIS Elements AR software (v.4.30.02).

### Senescence in young and old fibroblast cultures

We assessed senescence-associated β-galactosidase activity (SA-β-gal) in fibroblast cultures using a histochemical staining kit (Sigma-Aldrich, CS0030) according to the manufacturer's recommendations. The nuclei were stained with DAPI (Life Technologies). For determining the proportion of senescent cells, 5–10 images were randomly taken per sample and uploaded on ImageJ (v.1.46r). Senescence rate was calculated by dividing the number of SA-β-gal$^+$ cells with the total number of cells (DAPI stain).

### FACS and analysis of primary fibroblasts

We performed FACS analysis and sorting of THY1$^+$PDGFRα$^+$ and THY1$^-$PDGFRα$^+$ cells from primary fibroblast cultures at passage 3. FACS analysis was performed on an LSR II flow cytometer (BD Biosciences), and FACS sorting was performed on a BD FACS Aria II sorter, using a 100 μm nozzle. FACS data were analysed using FlowJo v.10.0.7. Gating was determined using fluorescence-minus-one controls for each colour used in each FACS experiment to ensure that positive populations were solely associated with the antibody for that specific marker (Extended Data Fig. 10). For FACS analysis of cultured cells, fibroblasts were stained with phycoerythrin-conjugated CD140a (BioLegend, 135905) and FITC (fluorescein isothiocyanate)-conjugated CD90.2 (BioLegend, 105305).

EdU incorporation in fibroblast cultures was assessed by FACS using the Click-iT EdU Plus FACS PacBlue Kit (Invitrogen, C10636) in accordance with the manufacturer's instructions. In brief, fibroblasts were incubated in medium containing EdU (10 μM) for 4 h. Cells were then dissociated and resuspended in FACS buffer (1% BSA in PBS). Cell surface markers were stained with phycoerythrin-conjugated CD140a (BioLegend, 135905) and FITC-conjugated CD90.2 (BioLegend, 105305). Cells were then fixed (4% paraformaldehyde in PBS) and permeabilized, followed by click reaction to detect EdU, according to the manufacturer's instructions.

### RT–qPCR on cultured fibroblasts

To assess expression of fibroblast subpopulation-specific genes, young and old fibroblast cultures at passage 3 were FACS-sorted into THY1$^+$PDGFRα$^+$ and THY1$^-$PDGFRα$^+$ fibroblasts (see above for details) and purified fibroblast subpopulations were expanded until passage 5–9. THY1$^+$PDGFRα$^+$ and THY1$^-$PDGFRα$^+$ cells were then plated at a density of 50,000 cells per well in a well of a 6-well plate. After 4 days, RNA was isolated from these fibroblast cultures and cDNA synthesis was performed as described in 'RT–qPCR on iPS cells and differentiated cells from embryoid bodies'. Comparisons were made between pairs from the same original culture. *Hprt1* was used as housekeeping gene for normalization. All primer sequences are listed in Supplementary Table 6a.

### Knockdown of the transcription factor EBF2

To test the functional implication of specific transcriptional regulators, we performed shRNA knockdown experiments. FACS-purified THY1$^+$PDGFRα$^+$ and THY1$^-$PDGFRα$^+$ fibroblasts at passages 4–9 were plated at a density of 50,000 cells per well in a 6-well plate. One day after plating, cells were infected by lentiviruses expressing shRNAs. Two independent lentiviral shRNA vectors against *Ebf2* were used (Sigma-Aldrich, TRCN0000081515 and TRCN0000081514). As control, a lentiviral shRNA vector against luciferase was used (Sigma-Aldrich, SHC007V). To produce lentiviruses, we followed the protocol described above (see 'Lentiviral production for reprogramming'). Viral supernatant was collected at 24 h after transfection, centrifuged at 3,000 r.p.m. for 15 min and transferred into a fresh tube. Next, 0.7 ml of the crude virus supernatant was added to THY1$^+$PDGFRα$^+$ and THY1$^-$PDGFRα$^+$ fibroblasts. The medium was changed 24 h after infection, and the cells were maintained in fibroblast growth medium for another 48 h before RNA collection. RNA collection and purification, and RT–qPCR, were performed as described above (see 'RT–qPCR on cultured fibroblasts').

### Overexpression of the transcription factor EBF2

Fibroblasts from young mice at passage 3 were plated at a density of 20,000 cells per well in a 12-well plate. One day after plating, cells were infected by lentiviruses expressing *Ebf2* or a vector control. To produce lentiviruses (see 'Lentiviral production for reprogramming'), the following vectors were used: 20 μg of pLenti-Ebf2-Myc-DDK (OriGene, MR224591L3) or 20 μg of pLenti-C-Myc-DDK (OriGene, PS100064), 12.6 μg of psPAX2 (Addgene 12260) and 3.7 μg of VSV-G. After 6 h of transfection, the medium was replaced by 7 ml fresh growth medium. At 24 and 48 h after transfection, viral supernatants were collected and centrifuged at 3,000 r.p.m. for 15 min and subsequently transferred into fresh tubes. Viral supernatant, collected from two 10-cm dishes of HEK293T cells at both time points, was concentrated by centrifugation at 16,500 r.p.m. for 1.5 h at 4 °C. The pellet was then resuspended in 2.5 ml fibroblast growth medium with polybrene (8 μg ml$^{-1}$, MilliporeSigma, TR1003G). Next, 0.35 ml of the concentrated virus was added to each well of fibroblasts. The medium was changed 24 h after infection, and the cells were maintained in fibroblast growth medium for another 24 h before RNA collection. RNA collection and purification, and RT–qPCR, were performed as described above (see 'RT–qPCR on cultured fibroblasts').

### Proliferation rate of young and old fibroblast cultures

Proliferation rate was assessed by plating young and old fibroblasts at a density of 50,000 cells per well of a 6-well plate in fibroblast growth medium. Every second day for up to 6 days, independent cultures were trypsinized and the number of cells in the cell suspension was counted manually using a haemocytometer. A growth slope was determined as the slope of the regression line based on the data points (cell numbers).

### FACS and analysis of fibroblasts in vivo in tissues

We isolated fibroblasts from the ears of young and old mice for FACS, quantification and transcriptomic analysis. In brief, ears were dissected from animals, cut into small fragments (around 1 mm$^2$) and digested in DMEM (Invitrogen, 11965-092) supplemented with 0.14 Wunsch units ml$^{-1}$ of Liberase DL (Roche, 5401160001) for 30 min at 37 °C. The fragments were washed with DMEM supplemented with 20% FBS (Gibco, 16000-044), funnelled through a 100-μm nylon mesh (Fisher Scientific, 08-771-19) and washed with fibroblast growth medium (DMEM supplemented with 10% FBS and 1% PSQ). A second filtering was performed using a 40-μm nylon mesh (Fisher Scientific, 08-771-1), followed by a washing step with fibroblast growth medium. Finally, cells were washed with FACS buffer (PBS, 1% BSA, 500 nM EDTA) and resuspended in FACS buffer to be stained for FACS analysis. FACS analysis and sorting was performed on a BD FACS Aria II sorter, using a 100-μm nozzle. FACS data were analysed using FlowJo v.10.0.7. Gating was determined using

fluorescence-minus-one controls for each colour used in each FACS experiment to ensure that positive populations were solely associated with the antibody for that specific marker (Extended Data Fig. 10). For in vivo FACS analysis and sorting the following antibodies were used: CD140a (BioLegend, 135905), CD90.2 (BioLegend, 105305), TER119 (Biolegend, 116234), CD326 (Thermo Fisher Scientific, 50-163-76), CD45 (Biolegend, 103126), CD31 (Biolegend, 102422), CD202b (Thermo Fisher Scientific, 15-5987-82), brilliant violet 421 streptavidin (Biolegend, 405226) and DAPI staining solution (Thermo Fisher Scientific, 62248).

### Bulk RNA-seq of THY1$^-$PDGFRα$^+$ and THY1$^+$PDGFRα$^+$ cells from the ears of young and old mice, before and after wounding

To determine whether cells could express cytokines in vivo, we profiled changes in transcriptomic features in fibroblast subpopulations in tissues from young and old mice, before and after wounding (see 'Wounding and wound healing experiments' for details on wounding experiments). RNA-seq was performed on freshly isolated THY1$^+$PDGFRα$^+$Lin$^-$ or THY1$^-$PDGFRα$^+$Lin$^-$ (defined as PDGFRα$^+$CD45$^-$CD31$^-$EpCAM$^-$TER119$^-$TIE2$^-$) (see above for isolation). Cells from 2–3 young or old mice were pooled together to obtain 500 cells of each population for each biological replicate. RNA isolation and generation of RNA-seq libraries were performed using the Clontech SmartSeq v.4 Ultra-Low Input RNA kit (Clontech). Cells were FACS-sorted directly into lysis buffer and cDNA was prepared as described by the manufacturer. Each cDNA library was analysed using a High Sensitivity chip on an Agilent 2100 Bioanalyzer. To generate sequencing libraries, 0.15 ng of each cDNA library was used as input for the Nextera XT kit, following the manufacturer's recommendations. Cells were indexed using the Nextera XT Index kit v.2 set A, and were subsequently multiplexed and sequenced on Illumina NextSeq-500 High Output Flow Cell (400 M), using 75-bp paired-end reads.

### Assessment of reprogramming efficiency of THY1$^+$PDGFRα$^+$ and THY1$^-$PDGFRα$^+$ fibroblasts

To determine the intrinsic reprogramming efficiency of THY1$^+$PDGFRα$^+$ and THY1$^-$PDGFRα$^+$ fibroblasts, cells were FACS-purified (see above) and plated at a density of 100,000 cells per well of a 6-well plate. Reprogramming was induced as described above (see 'Reprogramming of young and old fibroblasts to iPS cells and characterization of the iPS cells'). In these experiments, 0.1% gelatin-coated plates (Tribec Science, TBS8004) without feeders were used to avoid confounding factors from the feeder cells. Fibroblasts infected with lentiviruses expressing the OSKM factors were maintained on fibroblast growth medium until day 7 after replating and then switched to ES cell medium until days 12–13. Reprogramming efficiency was assessed by AP staining as described above (see 'Assessment of reprogramming efficiency'). This analysis revealed that activated THY1$^+$PDGFRα$^+$ fibroblasts intrinsically reprogram less efficiently compared to THY1$^-$PDGFRα$^+$ non-activated fibroblasts, and that old non-activated fibroblasts (THY1$^-$PDGFRα$^+$) also reprogrammed less efficiently than young THY1$^-$PDGFRα$^+$ fibroblasts (Extended Data Fig. 5q).

### Assessment of reprogramming efficiency of fibroblasts with swapped conditioned medium

To assess the contribution of extrinsic factors for reprogramming efficiency, we performed experiments swapping conditioned medium. For the conditioned medium experiments THY1$^+$PDGFRα$^+$ or THY1$^-$PDGFRα$^+$ fibroblasts (passages 5–9), or good or bad old fibroblasts (passage 3) were plated at a density of 350–400,000 cells per 10-cm plate or at a density of 0.5–1 million cells per 15-cm plate in fibroblast growth medium. In parallel, cells were plated at a density of 100,000 cells per well of a 6-well plate to induce reprogramming as described above (see 'Reprogramming of young and old fibroblasts to iPS cells and characterization of the iPS cells'). Starting from day 1 after infection with OSKM factors, conditioned medium was collected from 10-cm or 15-cm dishes from the indicated cultures that were growing in parallel, centrifuged at 10,000 r.c.f. for 10 min at room temperature and added every day to the recipient cells by replacing the medium. From day 7 after replating onwards, ES cell medium was made using the conditioned medium from fibroblast cultures as a base. Owing to the positive effect of conditioned medium on cellular reprogramming, reprogramming efficiency was assessed earlier than in other experiments, at days 9–10 after infection. Reprogramming efficiency was assessed by AP staining as described above (see 'Assessment of reprogramming efficiency'). For experiments using THY1$^+$PDGFRα$^+$ or THY1$^-$PDGFRα$^+$ fibroblasts, comparisons were made between pairs of THY1$^-$PDGFRα$^+$ and THY1$^+$PDGFRα$^+$ from the same original culture. For experiments using conditioned medium from THY1$^+$PDGFRα$^+$ or THY1$^-$PDGFRα$^+$ fibroblasts, conditioned medium was collected from the THY1$^-$PDGFRα$^+$ or THY1$^+$PDGFRα$^+$ fibroblasts from the same original culture, and comparisons were made between the effect of the different conditioned media on the specific populations.

### Non-viral reprogramming

To test whether variation in reprogramming efficiency could be owing to lentiviral infection, we used a non-viral reprogramming protocol. Non-viral reprogramming was induced using the piggyback transposon system containing the OSKM factors[49]. In brief, FACS-purified THY1$^+$PDGFRα$^+$ or THY1$^-$PDGFRα$^+$ cells (passages 4–6) were plated in a well of a 6-well plate at a density of 100,000 cells per well and transfected by the piggyback transposon vector using Lipofectamine 3000 (Life Technologies, 11668027), according to the manufacturer's instructions. Transfected fibroblasts were maintained on fibroblast growth medium until day 7, and then switched to ES cell medium until days 16–19. Reprogramming efficiency was calculated by counting the number of AP$^+$ colonies in the well.

### Effect of cytokines and blocking antibodies on reprogramming efficiency

To test how cytokines impact reprogramming efficiency, the following recombinant cytokines and blocking antibodies were purchased from R&D systems and used according to the manufacturer's recommendations: recombinant mouse IL-6 (R&D systems, 406-ML-025), recombinant mouse IL-1β (R&D systems, 401-ML-025), recombinant mouse IL-4 (R&D systems, 404-ML-050), recombinant mouse TNF (R&D systems, 410-NT-050), recombinant mouse VEGF (R&D systems, 493-MV-025), and normal polyclonal goat IgG (R&D systems, AB-108-C), goat polyclonal mouse anti-IL-6 blocking antibody (R&D systems, AB-406-NA) and goat polyclonal mouse anti-TNF blocking antibody (R&D systems, AB-410-NA). For all experiments, the recombinant cytokines were resuspended according to the manufacturer's instructions and used in culture medium at a final concentration of 10 ng ml$^{-1}$. For the blocking antibody experiments, the blocking antibodies were pre-incubated with the corresponding cytokine or conditioned medium for 60–90 min before treatment. The blocking antibodies (or control IgG) were used at a concentration of 8 μg ml$^{-1}$. For these experiments, young and old fibroblasts at passage 3 were plated at a density of 100,000 cells per well of a 6-well plate. To avoid confounding factors from the feeder cells, cells were plated on 0.1% gelatin-coated plates (Tribec Science, TBS8004). Reprogramming was induced as described above (see 'Reprogramming of young and old fibroblasts to iPS cells and characterization of the iPS cells'). Starting from day 1 after infection with OSKM factors, cells were treated with specific cytokines or with conditioned medium together with blocking antibodies. Reprogramming efficiency was calculated by counting the number of AP$^+$ or SSEA1$^+$ colonies in the wells as described above (see 'Assessment of reprogramming efficiency'). For the cytokine experiments, comparisons were made between treated and untreated cells originating from the same infected pool of cells and thus infection efficiency was not taken into account.

### Western blot analyses

To test whether the cytokines used in this study induce their cognate pathways in fibroblasts, we performed western blot analyses. Young and old fibroblasts at passage 3 were plated at a density of 100,000–150,000 cells in a 6-cm dish in fibroblast growth medium. After plating for 24 h, cells were treated with the indicated cytokines or antibodies for 30 min. Cells were then lysed directly in the culture plates using ice-cold RIPA buffer (50 mM Tris-HCL pH 7.5, 150 mM NaCl, 2 mM EDTA, 1% NP-40, 0.1% SDS supplemented with 1 mM aprotinin, PMSF and PhosphoStop (Pierce)), scraped and transferred to Eppendorf tubes. Following addition of sample buffer (0.0945 M Tris-HCl pH 6.8, 9.43% glycerol, 2.36% w/v SDS and 5% β-mercaptomethanol), samples were resolved on 10% SDS–PAGE gels, transferred onto nitrocellulose membranes and blotted, according to the following protocol[50], using the following antibodies: phosphorylated STAT3 (Tyr705) (Cell Signaling Technology, 9145), STAT3 (Invitrogen, 44-364G), phosphorylated STAT6 (Tyr641) (Cell Signaling Technology, 9361), STAT6 (Cell Signaling Technology, 5397), phosphorylated AKT (Ser473) (Cell Signaling Technology, 4060), AKT (Cell Signaling Technology, 4691), phosphorylated NF-κB (Ser536) (Cell Signaling Technology, 3033), NFκB (Cell Signaling Technology, 8242), phosphorylated JNK1 and JNK2 (Thr183 and Tyr185) (Invitrogen, 44-682G), JNK1 (Invitrogen, 44-690G) and β-actin (Novus Biologicals, NB600-501). Membranes were incubated with HRP-conjugated anti-mouse (Calbiochem, 401215) or anti-rabbit secondary (Calbiochem, 401393) antibodies and visualized using enhanced chemiluminescence detection reagent (Amersham ECL, GE Healthcare).

### Wounding and wound healing experiments

To assess the change in the wound healing ability of mice with ageing, young (3–4 months) and old (24–26 months) C57BL/6 male mice from the NIA colony were anaesthetized in standard fashion by inhalation of 1–4% of isoflurane[51]. The hair on the dorsal aspect of both ears was shaved and cleaned with a 70% ethanol solution. Symmetric full-thickness skin wounds were induced on both ears by first gently pressing a 4-mm punch biopsy onto the dorsum of the ear at its cartilaginous base. Sharp scissors were then used to dissect away the wheel of skin while leaving the underlying connective tissue, cartilage and anterior skin. No dressing was applied post-operatively, and the wounds were allowed to heal without further intervention. Wound healing was assessed by imaging (using a standard iPhone 8S camera) every other day for 20 days. Wound closure was analysed by comparing the relative wound size at a given time to the original size immediately after the operation, performed as previously described[51]. A wound was considered closed when it was re-epithelialized for more than 95% of its original size[52]. The rate of individual wound healing was determined using the average of the resultant measurements from both ears per mouse. FACS analysis assessing the percentage of activated THY1$^+$PDGFRα$^+$Lin$^-$ cells in ears of young and old mice, before and after wounding, revealed that the fibroblasts in the wounds were predominantly activated (THY1$^+$PDGFRα$^+$) fibroblasts (Extended Data Fig. 7d). In line with this finding, all three populations identified in the single-cell RNA-seq analysis of all live cells in the old wounds exhibited enrichment for different aspects of the activated fibroblast state (Extended Data Fig. 9e).

### Single-cell RNA-seq of fibroblasts from young versus old wounds using 10x Genomics Chromium

To evaluate changes in the fibroblast composition of wounds with age, we performed single-cell RNA-seq of all live PDGFRα$^+$Lin$^-$ cells in the wounded area from young and old mice, 7 days after wounding. We pooled cells from 10 young (3–4 months) or 10 old (24–26 months) male C57BL/6 mice from the NIA aged colony. FACS sorting was performed as described above. Cells were sorted into chilled fibroblast growth medium. Cells were then spun down at 300*g* for 5 min at 4 °C and resuspended in fibroblast growth medium at a concentration of 263 cells per μl. Young and old cells were loaded onto a 10x Genomics Chromium chip as per the manufacturer's recommendations. Reverse transcription and library preparation was performed using the 10x Genomics Single Cell v.2 kit following the 10x Genomics protocol. One library from young mice and one library from old mice were multiplexed and sequenced on one lane of Illumina NextSeq-500 High Output Flow Cell (400 M), using 75-bp paired-end reads.

### Single-cell RNA-seq of the entire wounds of fast- versus slow-healing old mice using 10x Genomics Chromium

To determine the differences in the composition of cells from old mice with different wound healing trajectories, we performed single-cell RNA-seq of all live cells in the entire wounds of two old mice with slow-healing trajectories and two old mice with fast-healing trajectories, 7 days after wounding. Mice were sedated and mice were perfused with 20 ml of PBS with heparin sodium salt (50 U ml$^{-1}$) (Sigma Aldrich) to remove the blood, and ears were immediately harvested. Wounds were dissected and processed as described above. Live/dead staining was performed using 1 μg ml$^{-1}$ propidium iodide (Biolegend). FACS sorting was performed on a BD FACS Aria Fusion sorter using a 100-μm nozzle. Cells were sorted into chilled fibroblast growth medium. Cells were then spun down at 300*g* for 5 min at 4 °C and resuspended in fibroblast growth medium at a concentration of 1,000–1,500 cells per μl. Cells were loaded onto a 10x Genomics Chromium chip as described above. Two libraries from 2 fast old mice and two libraries from 2 slow old mice were multiplexed and sequenced on one lane of Illumina Novaseq 6000 S2, using 101bp paired-end reads.

### Quality control of 10x Genomics single-cell RNA-seq

For mapping, sequences obtained from sequencing using the 10x Genomics single-cell RNA-seq platform were de-multiplexed using the Cell Ranger package from 10x Genomics and mapped to the mm10 transcriptome using the Cell Ranger package (10x Genomics). Cells were removed from subsequent analysis if they were expressing fewer than 500 unique genes or expressed more than 10% mitochondrial reads. Levels of mitochondrial reads and numbers of Unique molecular identifiers were similar between the young and old mice (Extended Data Fig. 8a) and between the old mice with different wound healing capacities (Extended Data Fig. 8i), indicating that there was no systematic bias in the libraries between the conditions tested. Average gene detection in each library was also similar between the conditions tested (Extended Data Fig. 8a, i). Our study includes 13,833 total cells, with 3,036 PDGFRα$^+$Lin$^-$ cells from wounds of pooled young and old mice (1,592 young cells and 1,444 old cells) and 10,797 cells from individual wounds from old mice with different wound healing capacities (fast old 1, 2,533 cells; fast old 2, 2,376 cells; slow old 1, 3,761 cells; slow old 2, 2,127 cells).

### *t*-SNE analysis of single-cell RNA-seq datasets and identification of cell clusters

To analyse the single-cell RNA-seq data, we performed *t*-SNE clustering using the Seurat R Package (v.2.3.4) with the first 30 principal components[53,54]. Identification of significant clusters was performed using the FindClusters() algorithm in the Seurat package, which uses a shared nearest neighbour (SNN) modularity optimization-based clustering algorithm[53,54]. Marker genes for each significant cluster were found using the Seurat function FindAllMarkers(). This analysis identified two main clusters of fibroblasts between young and old wounds and seven main clusters of cells between the old mice with different wound healing trajectories (Fig. 4b, c and Extended Data Fig. 8b, j). Cell types were determined using a combination of marker genes identified from the literature, PAGODA analysis and GO for cell types using the web-based tool Enrichr (http://amp.pharm.mssm.edu/Enrichr/). We note that some known components of the skin (for example, keratinocytes and epithelial cells) were not identified in these wounds, similar to a recent single-cell RNA-seq study on dorsal skin after wounding in young animals[14]. This

could be owing to wound composition, dissociation properties and survival during the FACS sorting protocol, as previous single-cell RNA-seq studies that identified epithelial cells in skin, have either specifically isolated epithelial cells using FACS[55] or used a different isolation protocol on unwounded skin[56–58].

### PAGODA on single-cell RNA-seq data from wounds of young and old mice or of old mice with different wound healing trajectories

We performed PAGODA analyses using raw counts for all genes that were considered expressed for analyses of individual datasets (Fig. 4d and Extended Data Figs. 8c, 9d) and using Seurat normalized counts for the combined analysis (Extended Data Fig. 9k). We performed three separate analyses: (1) young compared to old PDGFRα$^+$Lin$^-$ cells (Extended Data Fig. 8c); (2) cells identified as fibroblasts by Seurat from wounds of fast- compared to slow-healing old mice (Extended Data Fig. 9d); and (3) combined fibroblasts from both datasets (all PDGFRα$^+$Lin$^-$ cells together with the cell cluster identified as fibroblasts by Seurat) (Extended Data Fig. 9k). For gene sets, we used all KEGG pathways as well as the in vitro fibroblast ageing and fibroblast activation genes sets described above (see Supplementary Table 2b, f). We used the PAGODA pipeline with default parameters, unless stated otherwise, and used the SCDE package v.1.99.1 in R v.3.3. We did not account for cell cycle in these analyses. We noted that fibroblast subpopulation B did not contain cells from old/young in the combined analysis. This is probably owing to the fact that this subpopulation of fibroblast has some markers of the haematopoietic lineage (Extended Data Fig. 9d, k), and is probably depleted of the PFGDRα$^+$Lin$^-$ FACS-sorting technique that we used to isolate fibroblasts from the wounds of young and old mice.

### Violin plots for gene expression of single cells

To visualize the expression of individual genes, cells were grouped by cell type (as determined by PAGODA). The log-transformed and normalized gene expression values as calculated by Seurat were plotted for each cell as a violin plot with an overlying dot plot in R.

### Statistical analysis

For most experiments, young and old mice or samples were processed in an alternate manner rather than in two large groups, to minimize group effects. Although we did not do a bona fide power analysis, we took into account previous experiments to estimate the number of animals needed in each experiment. The exception is the wound healing experiment in Fig. 4, in which a power analysis was performed based on an initial experiment to determine the sample size required to detect a difference in variability with a 95% confidence interval. For all quantifications that were done with FACS or automated image quantification, no blinding was performed, including Fig. 3a, b and Extended Data Figs. 5b, 7d. The other experiments were blinded, with the exception of Fig. 1c and Extended Data Figs. 1l, m, 5e, 6e–h, k. Statistical analysis of the differences between age groups was performed using a unpaired two-tailed nonparametric Wilcoxon rank-sum test or a paired two-tailed or a one-tailed nonparametric Wilcoxon signed-rank test, unless otherwise stated. The statistical test applied was determined before performing the experiments. In cases in which the same recipient fibroblast culture was used in two independent experiments (Extended Data Figs. 5d, q, u, 6g), an average of the resultant measurements was determined. The nonparametric Fligner–Killeen test was used to test for differences in variance in reprogramming efficiency. *P* values were corrected for multiple hypothesis testing using Benjamini–Hochberg correction, unless otherwise stated, and were considered significant when $P < 0.05$.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

All raw sequencing reads for population RNA-seq, ChIP–seq and single-cell RNA-seq data can be found under BioProject PRJNA316110. The command and configuration files, in addition to a list of all versioned dependencies present in the running environment, are available on the Github repository for this paper (https://github.com/brunetlab/Mahmoudi_et_al_2018) (except for the code for the processing of metabolomics data, which is available upon request).


25. Ramírez-Solis, R. et al. Genomic DNA microextraction: a method to screen numerous samples. *Anal. Biochem.* **201**, 331–335 (1992).
26. Lo Sardo, V. et al. Influence of donor age on induced pluripotent stem cells. *Nat. Biotechnol.* **35**, 69–74 (2017).
27. Kareta, M. S. et al. Inhibition of pluripotency networks by the Rb tumor suppressor restricts reprogramming and tumorigenesis. *Cell Stem Cell* **16**, 39–50 (2015).
28. Vierbuchen, T. et al. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* **463**, 1035–1041 (2010).
29. Ackermann, M. & Strimmer, K. A general modular framework for gene set enrichment analysis. *BMC Bioinformatics* **10**, 47 (2009).
30. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
31. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
32. Hashimoto, T. B., Edwards, M. D. & Gifford, D. K. Universal count correction for high-throughput sequencing. *PLOS Comput. Biol.* **10**, e1003494 (2014).
33. Feng, J., Liu, T., Qin, B., Zhang, Y. & Liu, X. S. Identifying ChIP–seq enrichment using MACS. *Nat. Protocols* **7**, 1728–1740 (2012).
34. Zhang, Y. et al. Model-based analysis of ChIP–seq (MACS). *Genome Biol.* **9**, R137 (2008).
35. Ross-Innes, C. S. et al. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**, 389–393 (2012).
36. Li, Q., Brown, J. B., Huang, H. & Bickel, P. J. Measuring reproducibility of high-throughput experiments. *Ann. Appl. Stat.* **5**, 1752–1779 (2011).
37. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
38. Benayoun, B. A. et al. H3K4me3 breadth is linked to cell identity and transcriptional consistency. *Cell* **158**, 673–688 (2014).
39. Azuara, V. et al. Chromatin signatures of pluripotent cell lines. *Nat. Cell Biol.* **8**, 532–538 (2006).
40. Bernstein, B. E. et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315–326 (2006).
41. Guan, S., Price, J. C., Prusiner, S. B., Ghaemmaghami, S. & Burlingame, A. L. A data processing pipeline for mammalian proteome dynamics studies using stable isotope metabolic labeling. *Mol. Cell. Proteomics* **10**, M111.010728 (2011).
42. Smith, C. A., Want, E. J., O'Maille, G., Abagyan, R. & Siuzdak, G. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal. Chem.* **78**, 779–787 (2006).
43. van den Berg, R. A., Hoefsloot, H. C., Westerhuis, J. A., Smilde, A. K. & van der Werf, M. J. Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics* **7**, 142 (2006).
44. Xia, J., Sinelnikov, I. V., Han, B. & Wishart, D. S. MetaboAnalyst 3.0—making metabolomics more meaningful. *Nucleic Acids Res.* **43**, W251–W257 (2015).
45. Bronte, V. & Zanovello, P. Regulation of immune responses by l-arginine metabolism. *Nat. Rev. Immunol.* **5**, 641–654 (2005).
46. Phang, J. M., Liu, W., Hancock, C. & Christian, K. J. The proline regulatory axis and cancer. *Front. Oncol.* **2**, 60 (2012).
47. Rath, M., Müller, I., Kropf, P., Closs, E. I. & Munder, M. Metabolism via arginase or nitric oxide synthase: two competing arginine pathways in macrophages. *Front. Immunol.* **5**, 532 (2014).
48. Fan, J. et al. Characterizing transcriptional heterogeneity through pathway and gene set overdispersion analysis. *Nat. Methods* **13**, 241–244 (2016).
49. O'Malley, J. et al. High-resolution analysis with novel cell-surface markers identifies routes to iPS cells. *Nature* **499**, 88–91 (2013).
50. Schaffer, B. E. et al. Identification of AMPK phosphorylation sites reveals a network of proteins involved in cell invasion and facilitates large-scale substrate prediction. *Cell Metab.* **22**, 907–921 (2015).
51. Lam, M. T., Nauta, A., Meyer, N. P., Wu, J. C. & Longaker, M. T. Effective delivery of stem cells using an extracellular matrix patch results in increased cell survival and proliferation and reduced scarring in skin wound healing. *Tissue Eng. Part A* **19**, 738–747 (2013).
52. Barret, J. P. et al. Accelerated re-epithelialization of partial-thickness skin wounds by a topical betulin gel: results of a randomized phase III clinical trials program. *Burns* **43**, 1284–1294 (2017).
53. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).
54. Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902 (2019).
55. Joost, S. et al. Single-cell transcriptomics of traced epidermal and hair follicle stem cells reveals rapid adaptations during wound healing. *Cell Rep.* **25**, 585–597 (2018).
56. Cheng, J. B. et al. Transcriptional programming of normal and inflamed human epidermis at single-cell resolution. *Cell Rep.* **25**, 871–883 (2018).

57. Der, E. et al. Tubular cell and keratinocyte single-cell transcriptomics applied to lupus nephritis reveal type I IFN and fibrosis relevant pathways. *Nat. Immunol*. **20**, 915–927 (2019).
58. Sun, Z. et al. A Bayesian mixture model for clustering droplet-based single-cell transcriptomic data from population studies. *Nat. Commun*. **10**, 1649 (2019).
59. Mertens, J. et al. Directly reprogrammed human neurons retain aging-associated transcriptomic signatures and reveal age-related nucleocytoplasmic defects. *Cell Stem Cell* **17**, 705–718 (2015).
60. Luminex. *xPonent 3.1 Logistic Curve Fitting: Technical Notes*. https://www.luminexcorp.com/zh/download/xponent-3-1-logistic-curve-fitting-white-paper/ (2014).
61. Carter, A. C., Davis-Dusenbery, B. N., Koszka, K., Ichida, J. K. & Eggan, K. Nanog-independent reprogramming to iPSCs with canonical factors. Stem *Cell Rep*. **2**, 119–126 (2014).



**Acknowledgements** We thank D. Wagh from Stanford Functional Genomics Facility for help with 10x Genomics single-cell RNA-seq; C. Carswell-Crumpton and M. Weglarz from Stanford Shared FACS Facility for FACS support; L. Liu for help with shRNA experiments; V. Sebastiano and M. Kareta for help with iPS cell generation and quality assessment; V. Sebastiano, K. Andreasson, C. Weyand and T. Wyss-Coray for helpful discussions and input to the manuscript; L. Booth, T. Ruetz, P. P. Singh, X. Zhao, P. Navarro, B. Dulken and other members of the Brunet laboratory for helpful discussions and feedback on the manuscript; M. Buckley, B. Demmitt and P. P. Singh for helping with independent code checking; S. Chen, A. Freund, J. Goudeaux, M. Pesch, A. Roux and J. Reuter for feedback on the initial manuscript. The piggyback OSKM transposon constructs were a gift from K. Kaji. This work used the Genome Sequencing Service Center of the Stanford Center for Genomics and Personalized Medicine Sequencing Center, supported by the grant award NIH S10OD025212. Sorting was performed on instruments in the Stanford Shared FACS Facility obtained using the NIH grant award S10RR025518-01 and purchased by Parker Institute for Cancer Immunotherapy. Work was supported by NIH P01 GM099130 (to A.B., M.P.S. and J.W.), CIRM RB4-06087 (to A.B.), a generous philanthropic gift from M. and T. Barakett, the EMBO post-doctoral fellowship (to S.M.), the WennerGren post-doctoral fellowship (to S.M.), the Sweden-America Foundation post-doctoral fellowship (to S.M.), the Dean's Fellowship (to E.M.), K99/R00 AG049934 (to B.A.B.), the Stanford Graduate Fellowship (to L.X.), the NSF Graduate Research Fellowship (to L.X.), R00 AG049934 and the Hanson-Thorell family fellowship (to B.A.B.), the Hagey Laboratory for Pediatric Regenerative Medicine, The Gunn/Olivier fund, the Johnson Longaker fund and R01 GM116892 (to M.T.L.).


**Author contributions** L.X. and A.M contributed equally to this study. S.M., E.M. and A.B. planned the study, with input from M.P.S. and J.W. S.M. and E.M. performed the reprogramming experiments, and analysed and interpreted data. S.M., E.M. and L.X. wrote the manuscript with the help of A.B. S.M. generated, processed and analysed bulk and single-cell RNA-seq datasets, analysed the metabolomics data, and performed most THY1-related and conditioned medium experiments. E.M. generated and propagated transgene-free iPS cell lines. All other studies were done by both E.M and S.M., unless otherwise noted. A.M. and S.M. performed wound healing experiments under the supervision of M.T.L. L.X. helped with reprogramming, FACS, and immunofluorescence experiments. F.J. generated the in vitro single-cell RNA-seq data under the supervision of M.P.S. R.S. generated the ChIP–seq libraries under the supervision of J.W. K.H. helped with statistics and PAGODA analysis. X.L. performed metabolomics experiments and helped with metabolomics data analysis and validation under the supervision of M.P.S. K.D. helped with reprogramming and western blotting experiments. L.P. helped with reprogramming and RT–qPCR experiments. C.E.A. and Y.S. performed the induced neuron reprogramming experiment under the supervision of M.W. B.A.B. helped with analysis of the epigenomic data. A.L.S.C. identified and collected the human samples. All authors discussed the results and commented on the manuscript.

**Competing interests** The authors declare no competing interests.

**Additional information**

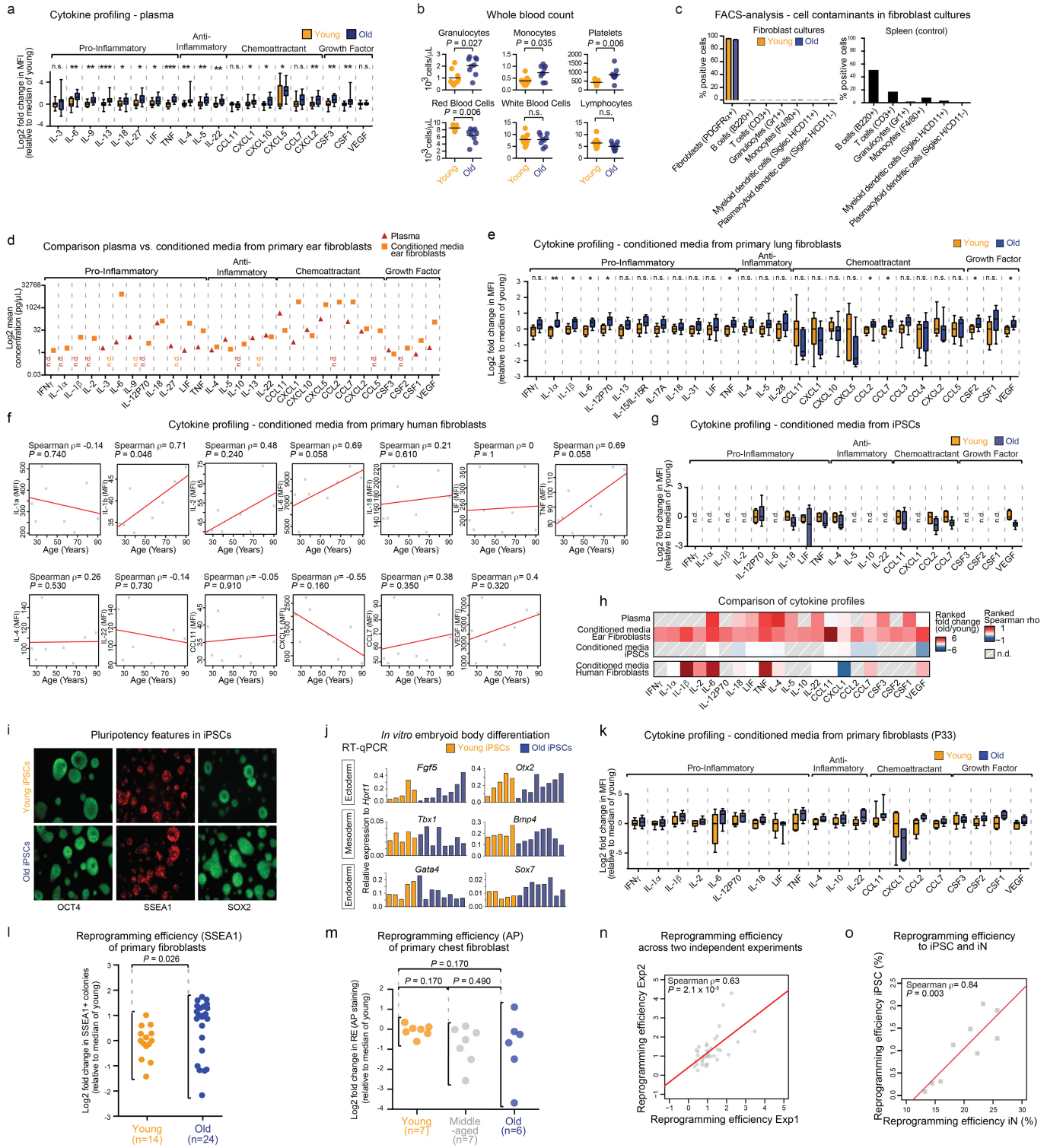

**Supplementary information** is available for this paper at https://doi.org/10.1038/s41586-019-1658-5.



**Correspondence and requests for materials** should be addressed to A.B.

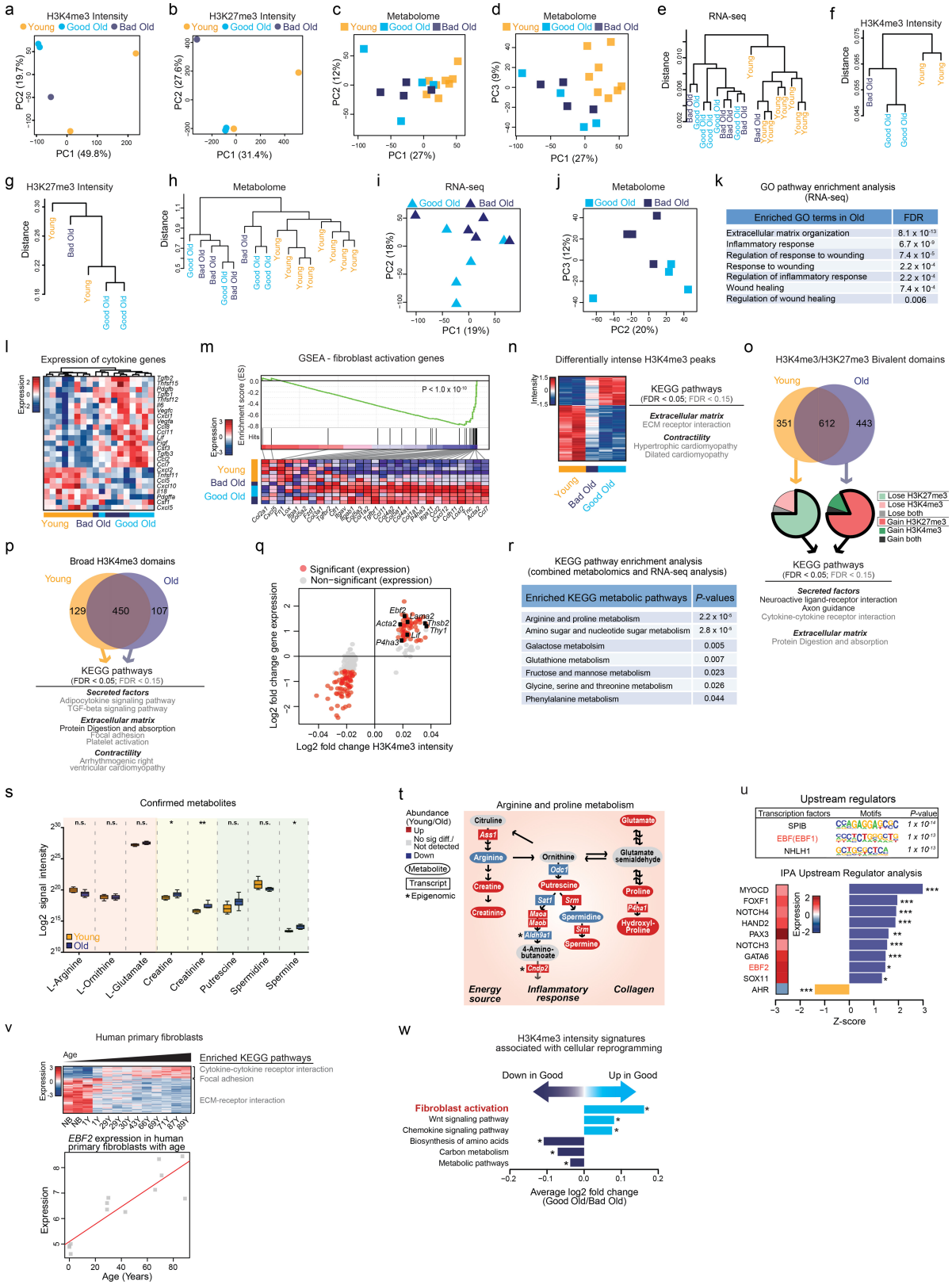


**Peer review information** *Nature* thanks Ellen Pure, Fuchou Tang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



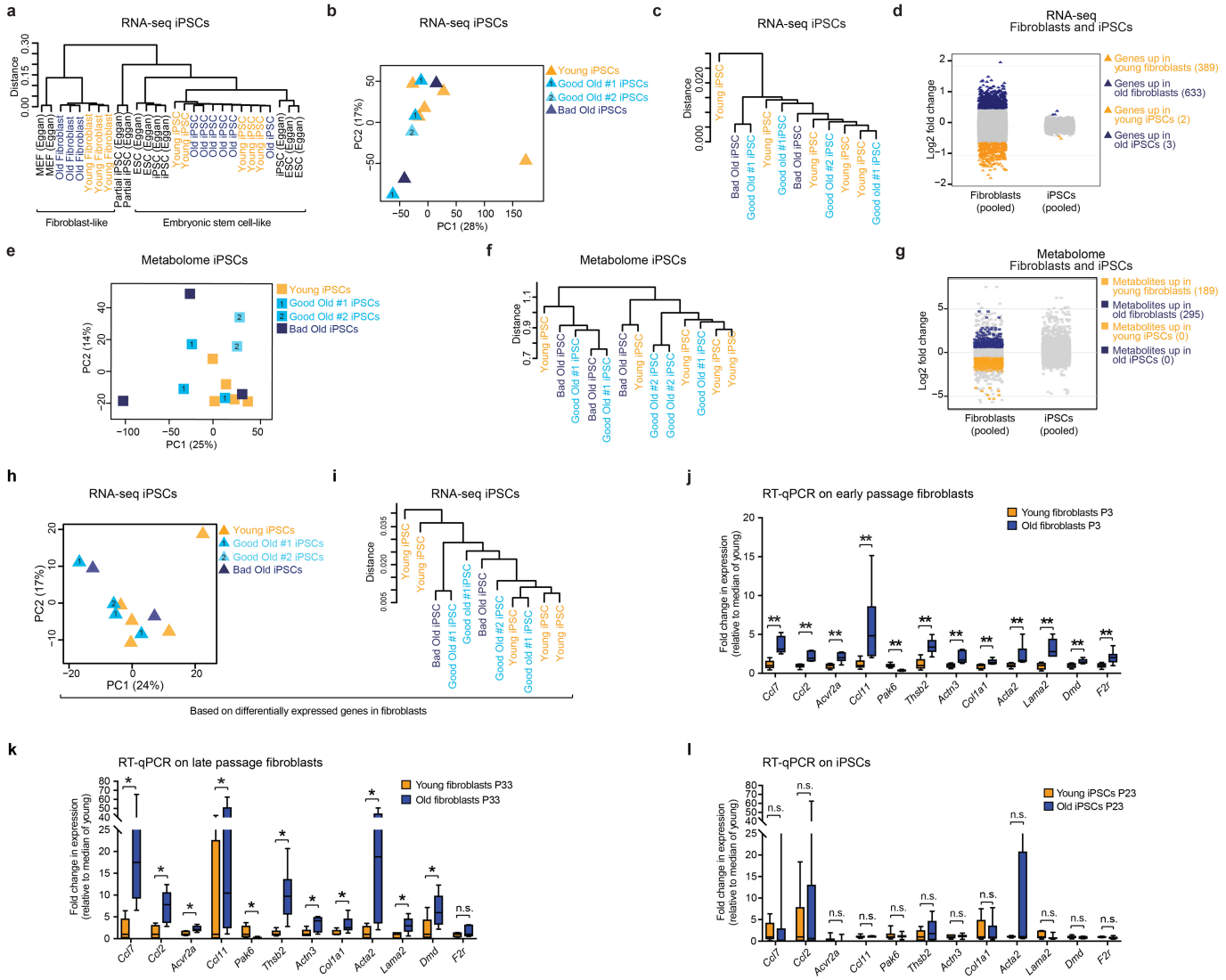

**Extended Data Fig. 1** | See next page for caption.

**Extended Data Fig. 1 | Primary old fibroblasts from mouse ear, mouse lungs and human skin secrete high levels of inflammatory cytokines, and the ability of individual cultures from ear fibroblasts to reprogram is stereotypical. a**, Cytokine profiling of plasma from young (3 months, $n = 21$) and old (29 months, $n = 19$) mice using Luminex multiplex cytokine assay (2 independent experiments). Box-and-whisker plot of $\log_2$-transformed fold change in MFI over median of young fibroblasts. Box plots depict the median and interquartile range, with whiskers indicating minimum and maximum values. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, n.s., not significant; two-tailed Wilcoxon rank-sum test with Benjamini–Hochberg correction. Exact $P$ values are in Supplementary Table 1a. **b**, Blood cell composition analysis of plasma from young (3 months, $n = 9$) and old (29 months, $n = 9$) mice using the Hemavet Multispecies Hematology Analyzer (2 independent experiments). Data are cell numbers per µl in whole-blood samples. Each dot represents cells from one mouse. Lines depict median. $P$ values, two-tailed Wilcoxon rank-sum test with Benjamini–Hochberg correction. **c**, Percentage of fibroblasts (PDGFRα$^+$) and immune cells in young (3 months, $n = 8$) and old (29 months, $n = 8$) fibroblast cultures at passage 3 (1 experiment), as determined by FACS using the indicated cell-type-specific surface markers. Data are mean ± s.e.m. Primary splenocytes from an 8-week-old mouse were used as a positive control (right). **d**, Comparison between cytokine profiles of plasma (red triangles) and conditioned medium from cultured ear fibroblasts at passage 3 (orange squares). Results are the mean $\log_2$-transformed concentrations (pg µl$^{-1}$) of cytokines detected. For exact concentrations, see Supplementary Table 1a, b. n.d., not detected. **e**, Cytokine profiles of conditioned medium from primary cultures (passage 3) of lung fibroblasts from young (3 months, $n = 8$) and old (20–24 months, $n = 9$) mice (2 independent experiments). Box-and-whisker plot of $\log_2$-transformed fold change in MFI over median of young fibroblasts. $^*P < 0.05$, $^{**}P < 0.01$; two-tailed Wilcoxon rank-sum test with Benjamini–Hochberg correction. Box plots as in **a**. Exact $P$ values are in Supplementary Table 1c. **f**, Cytokine profiles of conditioned medium from primary fibroblast cultures isolated from punch biopsy of pre-auricular skin of healthy human subjects of different ages. Results are shown as Spearman's rank correlation coefficient ($\rho$) between donor age (years) and cytokine levels (MFI) in human fibroblast cultures ($n = 8$) at passage 3 (1 experiment). Each dot represents cells from one individual. $P$ values, two-sided algorithm AS 89 in R. For multiple hypothesis testing, see Supplementary Table 1d. **g**, Cytokine profiles of conditioned medium collected from passage 23 cultures of iPS cell lines derived from young (3 months, $n = 4$) and old (29 months, $n = 6$) fibroblasts (1 experiment). Box-and-whisker plot of $\log_2$-transformed fold change in MFI over median of young iPS cells. Only cytokines that were detected at significantly different levels in young and old fibroblasts are shown (for a complete cytokine list and more details, see Supplementary Table 1e, f). Each dot represents an individual iPS cell line. $P$ values, one-tailed Wilcoxon rank-sum test with Benjamini–Hochberg correction. Exact $P$ values are in Supplementary Table 1e. Box plots as in **a**. **h**, Comparison of age-dependent changes in cytokine levels between plasma-incubated (described in **a**) and conditioned-medium-incubated mouse fibroblasts (described in Fig. 1b), their derived iPS cells (described in **g**) and from human fibroblasts (described in **f**) (Supplementary Table 1a, b, d, e) based on cytokines that are significantly different in conditioned medium from fibroblasts (Fig. 1b, bottom). Top (also presented in Fig. 1b), ranked fold change (old/young) in levels of the indicated cytokines in plasma, conditioned medium from mouse primary fibroblasts and iPS cells. Bottom, ranked Spearman $\rho$ correlations for the indicated cytokines in conditioned medium from human primary fibroblasts (see **f** for individual $\rho$ values). **i**,**j**, iPS cell lines derived from young and old mice show typical morphologies of mouse iPS cells, express similar levels of pluripotency markers and can give rise to cell types from all three germ layers upon embryoid body formation. **i**, Representative immunofluorescence images of iPS cell lines derived from young and old mice at passage 23, stained with the indicated antibodies (1 experiment). **j**, RT–qPCR on the indicated genes in embryoid bodies differentiated in vitro from iPS cell lines from young ($n = 5$) and old ($n = 8$) mice at passage 23 (1 experiment). Expression is presented as expression relative to the housekeeping gene *Hprt1*. Each bar represents one iPS cell line. **k**, Cytokine profiles of conditioned medium collected from cultures of young ($n = 7$) and old ($n = 7$) ear fibroblasts at passage 33 (1 experiment). Box-and-whisker plot of $\log_2$-transformed fold change in MFI over median of young fibroblasts. Only the cytokines that exhibited a significant difference in expression levels in conditioned medium from young and old fibroblasts at passage 3 are shown. $P$ values, one-tailed Wilcoxon rank-sum test with Benjamini–Hochberg correction. For a complete cytokine list and exact $P$ values, see Supplementary Table 2r. Box plots as in **a**. Note that the experiments in fibroblasts at passage 3 and 33 were conducted independently, and therefore statistical comparisons were restricted to within experiments. However, a direct comparison between the levels of secreted factors at passage 3 to 33 revealed that the levels of most cytokines decrease upon passaging. **l**, Reprogramming efficiency, assessed by SSEA1 staining, of young ($n = 14$) and old ($n = 24$) ear fibroblast cultures at passage 3 (3 independent experiments), as $\log_2$-transformed fold change over the median of young mice. Each dot represents a fibroblast culture from one mouse. $P$ value, Fligner–Killeen test to assess differences in variance between age groups. **m**, Reprogramming efficiency assessed by AP staining of young (3 months, $n = 7$), middle-aged (12–13 months, $n = 7$) and old (28–30 months, $n = 8$) chest fibroblast cultures at passage 3 (2 independent experiments), as $\log_2$-transformed fold change over the median of young mice. Dots as in **l**. $P$ values, Fligner-Killeen test to assess differences in variance between age groups with Benjamini–Hochberg correction. **n**, Reprogramming efficiency of fibroblast cultures are mainly stereotypical to fibroblast cultures from an individual mouse. Correlation plot depicting the reprogramming efficiency (assessed by AP staining) of fibroblast cultures reprogrammed in two experiments (separated by more than one month), with data from experiment 1 on the $x$ axis and data from experiment 2 on the $y$ axis. Data shown are from young (3 months, $n = 14$), middle-aged (12 months, $n = 6$) and old (29 months, $n = 18$) mice. Dots as in **l**. $P$ values, two-sided algorithm AS 89 in R. There was a positive correlation (Spearman rank correlation, $\rho = 0.63$, $P = 2.1 \times 10^{-5}$) between reprogramming efficiencies of fibroblast cultures from the same mouse. **o**, Correlation plot depicting the ability of young ($n = 4$) and old ($n = 5$) fibroblast cultures at passage 3 to reprogram into neurons (iN, assessed by PSA-NCAM) or to iPS cells (iPSC, assessed by AP) (1 experiment). Dots as in **l**. $P$ values, two-sided algorithm AS 89 in R. There was a significant positive correlation (Spearman rank correlation, $\rho = 0.84$, $P = 0.003$) between these two features. For individual experiments in **a**, **b**, **e**, **l**–**n**, see Supplementary Table 7.



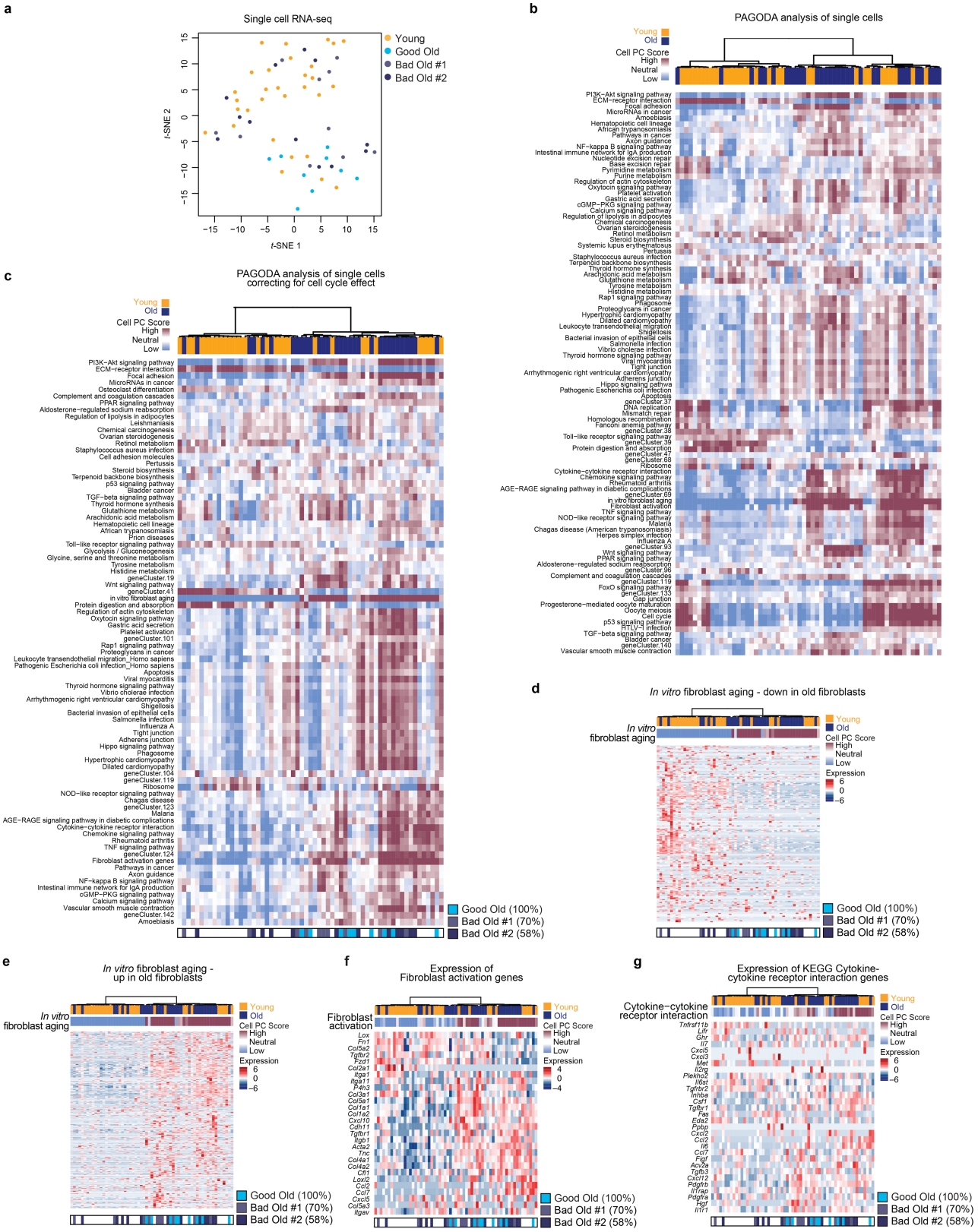

**Extended Data Fig. 2 |** See next page for caption.

**Extended Data Fig. 2 | Old fibroblasts exhibit distinct transcriptomic, epigenomic and metabolomics profiles compared to young fibroblasts.** **a**, **b**, PCA of H3K4me3 (**a**) and H3K27me3 (**b**) peak intensities from young ear fibroblast cultures (3 months, $n = 2$), old cultures that reprogram well (good old, 29 months, $n = 2$) and old cultures that reprogram poorly (bad old, 29 months, $n = 1$) (1 experiment). Principal components (PCs) 1 and 2 are shown. **c**, PCA of metabolomes of young ($n = 8$), good old ($n = 4$) and bad old ($n = 4$) ear fibroblast cultures (1 experiment). Principal components 1 and 2 are shown. **d**, PCA of metabolomes of young ($n = 8$), good old ($n = 4$) and bad old ($n = 4$) ear fibroblast cultures (1 experiment). Principal components 1 and 3 are shown. **e**, Unsupervised hierarchical clustering of transcriptomes (RNA-seq) of young ($n = 8$), good old ($n = 5$) and bad old ($n = 5$) ear fibroblast cultures (3 independent experiments). Hierarchical clustering was performed using correlation-based dissimilarity (Pearson's) as distance measure and average for linkage analysis. The *y* axis indicates the similarity between samples. **b**–**e**, Mice were the same ages as in **a**. **f**–**h**, Unsupervised hierarchical clustering of H3K4me3 (**f**) and H3K27me3 (**g**) peaks and metabolomes (**h**) described in **a**, **b** and **c**, respectively. Hierarchical clustering was performed as in **e**. **i**, PCA of transcriptomes (RNA-seq) of good and bad old ear fibroblast cultures. Principal components 1 and 2 are shown. **j**, PCA of metabolomes of good old ($n = 4$) and bad old ($n = 4$) ear fibroblast cultures. Principal components 2 and 3 are shown. **k**, Selected GO terms enriched in the old transcriptomes (young, $n = 8$; old, $n = 10$), with corresponding FDR-adjusted *P* values (one-sided Fisher's exact test with Benjamini–Hochberg correction). For a complete list, see Supplementary Table 2d. **l**, Heat map showing expression (VST-transformed read counts, scaled row-wise) of selected cytokine genes. The scale for expression fold changes is indicated on the left. **m**, Gene set enrichment analysis (GSEA) plot depicting the transcriptomes of young ($n = 8$) and old ($n = 10$) fibroblasts. Top, genes associated with fibroblast activation (see Supplementary Table 2f) are enriched in the old fibroblasts ($P < 1.0 \times 10^{-10}$, two-sided nominal *P* value). Bottom, heat map shows expression of fibroblast activation genes (VST-transformed read counts, scaled row-wise). The scale for expression fold changes is indicated on the left. **n**–**q**, Analysis of the epigenomic data as described in **a**, **b**. In line with the transcriptomic data, age-dependent changes in epigenomic landscape also revealed enrichment of pathways involved in fibroblast activation, such as cytokines, extracellular matrix components and contractility-related features. **n**, Left, heat map shows the H3K4me3 peaks within promoter regions that exhibit a significant difference in intensity with age, assessed by Diffbind. Peak intensity is shown as VST-transformed read counts, scaled row-wise. The scale for peak intensity fold changes is indicated on the left. Right, selected enriched KEGG pathways colour coded according to significance (one-sided Fisher's exact test with Benjamini–Hochberg correction; black, FDR-adjusted $P < 0.05$; grey, FDR-adjusted $P < 0.15$). For a complete list of significant KEGG terms, see Supplementary Table 2h. **o**, Top Venn diagram depicts the overlap of bivalent domains within promoters regions of young and old fibroblasts. For details, see 'Chromatin immunoprecipitation followed by sequencing and analysis of the epigenomic landscape'. Middle, pie charts show how the unique bivalent domains in young fibroblasts change in old fibroblasts (left pie chart) and vice versa (right pie chart). Bottom, selected enriched KEGG pathways colour coded according to significance. For a complete list of KEGG terms, see Supplementary Table 2l. **p**, Top, Venn diagram depicting the overlap of broad H3K4me3 domains within promoters regions of young and old fibroblasts. For details see 'Chromatin immunoprecipitation followed by sequencing and analysis of the epigenomic landscape'). Bottom, selected enriched KEGG pathways colour coded according to significance. For a complete list of significant KEGG terms, see Supplementary Table 2j. **q**, The relationship between changes in H3K4me3 peak intensity (as described in **a**) and gene expression for H3K4me3 peaks that are significantly different between the age groups (as described in Fig. 2b). The analysis was restricted to H3K4me3 peaks that are within promoter regions, defined as the transcriptional start site ±2 kb. The *y* axis denotes the $\log_2$-transformed fold change in gene expression between young and old fibroblasts for the gene, and *x* axis denotes the $\log_2$-transformed fold change of H3K4me3 peak intensity assigned to the gene. Genes of interest are labelled. **r**–**t**, Pathway analysis of all putatively identified metabolites that were significantly different between young and old fibroblasts (described in **c**), as well as the differentially expressed genes (described in Fig. 2b; FDR-adjusted $P < 0.05$, absolute fold change >1.5), using the MetaboAnalyst tool[44]. Note that the MetaboAnalyst tool does not provide multiple-hypothesis-corrected *P* values (one-sided Fisher's exact test). **s**, Box plots showing the $\log_2$-transformed signal intensities of selected metabolites in the arginine and proline pathway from the metabolic profiling of young and old fibroblasts cultures at passage 3 (as described in **c**), for which the identity was confirmed using commercially available standards. Box plots depict the median and interquartile range, with whiskers indicating minimum and maximum values. $^*P < 0.05$, $^{**}P < 0.01$; two-tailed Wilcoxon rank-sum test with *q* value correction; L-arginine, $P = 0.069$; L-ornithine, $P = 0.690$; L-glutamate, $P = 0.055$; creatine, $P = 0.022$; creatinine, $P = 0.002$; putrescine, $P = 0.081$; spermidine $P = 0.094$; spermine, $P = 0.016$. **t**, Schematic representation of the biological functions of key metabolites and genes in the arginine and proline metabolic pathway, and how they relate to regulation of inflammatory cytokines and extracellular matrix synthesis[45–47]. Abundance of putative metabolites (oval) and gene transcripts (squares) in old fibroblasts is colour coded (red, higher in old; blue, lower in old; grey, not significantly different or not detected). Epigenomic changes are indicated with black asterisks. **u**, Top, top 3 motifs found in promoters of differentially expressed genes between young and old fibroblasts described in Fig. 2b, using HOMER motif analysis. Bottom, top 10 putative upstream regulators identified by the IPA database that are differentially expressed between young and old fibroblasts (FDR-adjusted $P < 0.05$, absolute fold change >1.5). Heat map depicts $\log_2$-transformed fold change in expression (old/young) calculated using DESeq2. The transcription factor identified across both analyses (EBF2) is in red. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$; cumulative hypergeometric distribution (Homer motif analysis), activation *z*‐score in IPA (upstream regulator analysis). For a complete list of significant motifs and upstream regulators, and exact *P* values, see Supplementary Table 2n. **v**, Top, heat map of differentially expressed genes in a regression analysis from young to old healthy human primary fibroblasts[59] ($n = 13$, FDR-adjusted $P < 0.05$). Expression is shown as VST-transformed read counts, scaled row-wise. The depicted KEGG pathways are FDR-adjusted $P < 0.15$ (one-sided Fisher's exact test with Benjamini–Hochberg correction) (Supplementary Table 2o, p). Bottom, VST-transformed expression of *EBF2* across human samples as a function of age (years). Each square represents transcripts from a patient. NB, newborn. **w**, Pathway enrichment analysis of KEGG pathways associated with enhanced (up) or reduced (down) reprogramming, comparing H3K4me3 peak intensities between the top ($n = 1$) and the bottom ($n = 1$) old reprogramming cultures in our datasets (see Supplementary Table 2a). All depicted KEGG pathways were significantly enriched (FDR-adjusted $P < 0.05$). $^*P < 0.05$; two-sided nominal *P* value with Benjamini–Hochberg correction. For a complete list of KEGG terms and exact *P* values, see Supplementary Table 3f.



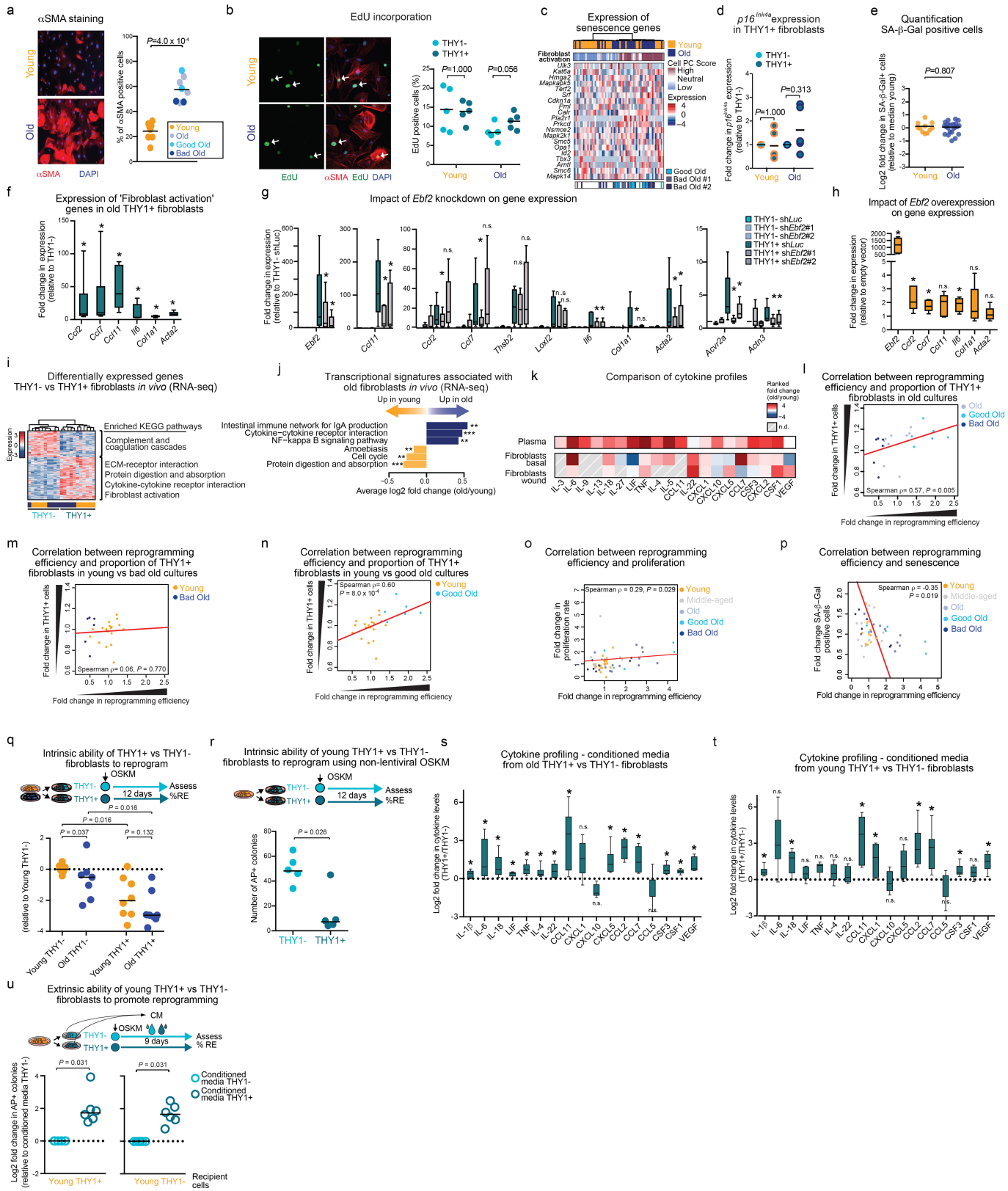

**Extended Data Fig. 3** | See next page for caption.

**Extended Data Fig. 3 | Reprogramming erases features of inflammageing and variability between mice. a–l**, To test whether transcriptomic and metabolomics features of inflammageing could be erased by reprogramming, iPS cell lines from young and old fibroblasts at passage 23 were profiled for their transcriptome (RNA-seq) and metabolome. **a**, All iPS cell lines ($n = 11$) cluster with previously established bona fide iPS cell lines ($n = 3$) and ES cells ($n = 4$)[61]. Unsupervised hierarchical clustering based on overall transcriptomes of the indicated cell types. The hierarchical clustering was performed using correlation-based dissimilarity (Pearson's) as distance measure and average for linkage analysis. The *y* axis indicates the similarity between samples. **b**, PCA of whole transcriptomes from RNA-seq data of iPS cell lines derived from young (3 months, $n = 5$) and old (29 months, $n = 6$) mice at passage 23. Principal components (PCs) 1 and 2 are shown. **c**, Unsupervised hierarchical clustering of transcriptomes described in **b**. Hierarchical clustering was performed as in **a**. **d**, Strip plot illustrating the $\log_2$-transformed fold expression changes of all genes with age for fibroblasts (left; described in Fig. 2b) and iPS cells (right; described in **b**). Genes detected as significantly upregulated or downregulated (DESeq2, FDR-adjusted $P < 0.05$, absolute fold change >1.5) with age, are shown in blue and yellow, respectively (see Supplementary Table 2q). **e**, PCA of metabolomes of iPS cell lines derived from young ($n = 5$) and old ($n = 8$) mice at passage 23. Ages as in **b**. Untargeted metabolomics profiles were generated using ultra-high performance liquid chromatography–mass spectrometry. Principal components 1 and 2 are shown. **f**, Unsupervised hierarchical clustering of metabolomics profiles described in **e**. Hierarchical clustering was performed as in **a**. **g**, Strip plot illustrating the $\log_2$-transformed fold change in signal intensity of all metabolic features with age for fibroblasts (left; described in Extended Data Fig. 2c) and iPS cells (right; described in **e**). Metabolic features detected as significantly up or down (using a two-tailed Wilcoxon rank-sum test with *q* value correction) with age are shown in blue and yellow, respectively. **h**, **i**, PCA (**h**) and unsupervised clustering (**i**) of iPS cell lines derived from young ($n = 5$) and old ($n = 6$) mice at passage 23, based on solely the genes that were significantly differentially expressed between young and old at fibroblast level. Ages as in **b**. Principal components 1 and 2 are shown. Hierarchical clustering was performed as in **a**. **j–l**, RT–qPCR of the indicated genes in fibroblasts cultures at passage 3 (**j**) and 33 (**k**), and iPS cell cultures at passage 23 (**l**). The genes shown represent the three major groups of features associated with fibroblast activation that change with age in fibroblasts. Box-and-whisker plot of $\log_2$-transformed fold change in MFI over median of young fibroblasts. Box plots depict the median and interquartile range, with whiskers indicating minimum and maximum values. Data are from young ($n = 6$) and old ($n = 6$) fibroblast cultures at passage 3, young ($n = 5$) and old ($n = 6$) fibroblast cultures at passage 33, young ($n = 6$) and old ($n = 7$) iPS cell cultures at passage 23. Ages as in **b**. $^*P < 0.05$, $^{**}P < 0.01$; one-tailed Wilcoxon rank-sum test with Benjamini–Hochberg correction. For **j**: *Ccl7* (also known as *Mcp3*), $P = 0.004$; *Ccl2* (also known as *Mcp1*), $P = 0.004$; *Acvr2a* (which encodes ACVR2α), $P = 0.006$; *Ccl11* (also known as *Eotaxin*), $P = 0.004$; *Pak6*, $P = 0.004$; *Thsb2*, $P = 0.004$; *Actn3*, $P = 0.006$; *Col1a1* (which encodes COL1α1), $P = 0.004$; *Acta2* (which encodes αSMA), $P = 0.004$; *Lama2*, $P = 0.004$; *Dmd*, $P = 0.008$; *F2r*, $P = 0.008$. For **k**: *Ccl7*, $P = 0.027$; *Ccl2*, $P = 0.027$; *Acvr2a*, $P = 0.027$; *Ccl11*, $P = 0.049$; *Pak6*, $P = 0.027$; *Thsb2*, $P = 0.026$; *Actn3*, $P = 0.026$; *Col1a1*, $P = 0.035$; *Acta2*, $P = 0.027$; *Lama2*, $P = 0.027$; *Dmd*, $P = 0.027$; *F2r*, $P = 0.229$. For **l**: *Ccl7*, $P = 0.800$; *Ccl2*, $P = 0.800$; *Acvr2a*, $P = 0.800$; *Ccl11*, $P = 1.000$; *Pak6*, $P = 0.800$; *Thsb2*, $P = 1.000$; *Actn3*, $P = 0.800$; *Col1a1*, $P = 1.000$; *Acta2*, $P = 1.000$; *Lama2*, $P = 1.000$; *Dmd*, $P = 1.000$; *F2r*, $P = 1.000$. Note that the experiments were conducted independently in fibroblasts at passage 3, 33 and iPS cells, and therefore the statistical comparisons indicated were restricted to each independent experiment. However, a comparison between the expression of secreted factors at passage 3 to 33 shows that expression of *Ccl11*, but not *Ccl2* and *Ccl7*, significantly decreases upon passaging.



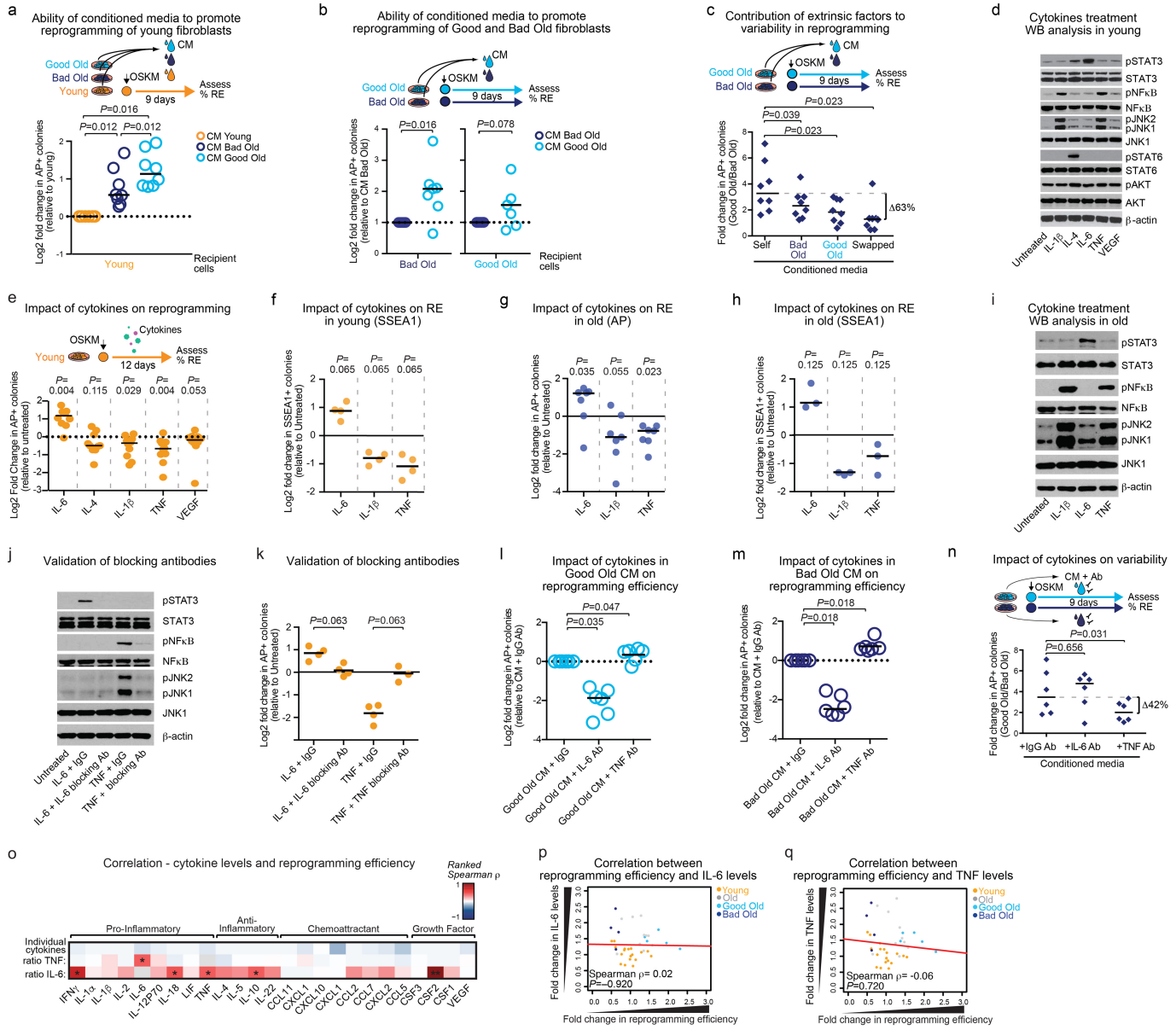

**Extended Data Fig. 4** | See next page for caption.

**Extended Data Fig. 4 | Correlation between the fibroblast activation signature and reprogramming efficiency in single-cell RNA-seq data.** **a**, Single-cell RNA-seq of young ($n = 30$ cells), good and bad old ($n = 31$ cells) fibroblast cultures. *t*-SNE was performed on the VST-transformed read counts of all detected genes (analysed using DESeq2). Each dot represents a single fibroblast transcriptome. **b**, PAGODA of single-cell RNA-seq data from young and old fibroblasts at passage 3 performed using raw expression counts and all KEGG pathways, the in vitro fibroblast ageing, the fibroblast activation and de novo gene sets. Hierarchical clustering is based on 97 significantly overdispersed gene sets and the 405 genes driving the significantly overdispersed gene sets. Top, heat map of single cells from young and old fibroblast cultures. Middle, heat map of the separation of cells based on their principal component scores for the significantly overdispersed gene sets. Top heat map, PAGODA clustering of cells. Maroon and blue colours indicate increased and decreased expression of the associated gene sets, respectively. **c**, PAGODA as described in Extended Data Fig. 4c. Top, heat map of single cells from young and old fibroblast cultures. Middle, heat map of separation of cells based on their principal component scores for the significantly overdispersed gene sets. Top heat map, PAGODA clustering of cells. Maroon and blue colours indicate increased and decreased expression of the associated gene sets, respectively. **d**, PAGODA as described in **c**. Middle, heat map of separation of cells based on their principal component scores for the in vitro fibroblast ageing signature. Bottom, heat map of the expression of the genes that are part of the in vitro fibroblast ageing signature, and decrease with age; expression is shown as VST-transformed read counts, scaled row-wise. The scale for expression fold changes is indicated on the right. The bottom heat map indicates the cells that originate from good and bad old cultures. **e**, PAGODA as described in **c**. Middle, heat map of separation of cells based on their principal component scores for the in vitro fibroblast ageing signature. Bottom, heat map of expression of the genes that are part of the in vitro fibroblast ageing signature, and increase with age; expression is shown as VST-transformed read counts, scaled row-wise. The scale for expression fold changes is indicated on the right. The bottom heat map indicates the cells that originate from good and bad old cultures. **f**, PAGODA as described in **c**. Middle, heat map of the separation of cells based on their principal component scores for the fibroblast activation signature. Bottom, heat map of the expression of the genes that are part of the fibroblast activation gene set; expression is shown as VST-transformed read counts, scaled row-wise. The scale for expression fold changes is indicated on the right. The bottom heat map indicates the cells that originate from good and bad old cultures. **g**, PAGODA as described in **c**. Middle, heat map of the separation of cells based on their principal component scores for the KEGG cytokine–cytokine receptor interaction gene set. The heat map shows expression of the top 30 overdispersed genes in the KEGG cytokine–cytokine receptor interaction pathway; expression is shown as VST-transformed read counts, scaled row-wise. The scale for expression fold changes is indicated on the right. The bottom heat map indicates the cells that originate from good and bad old cultures.



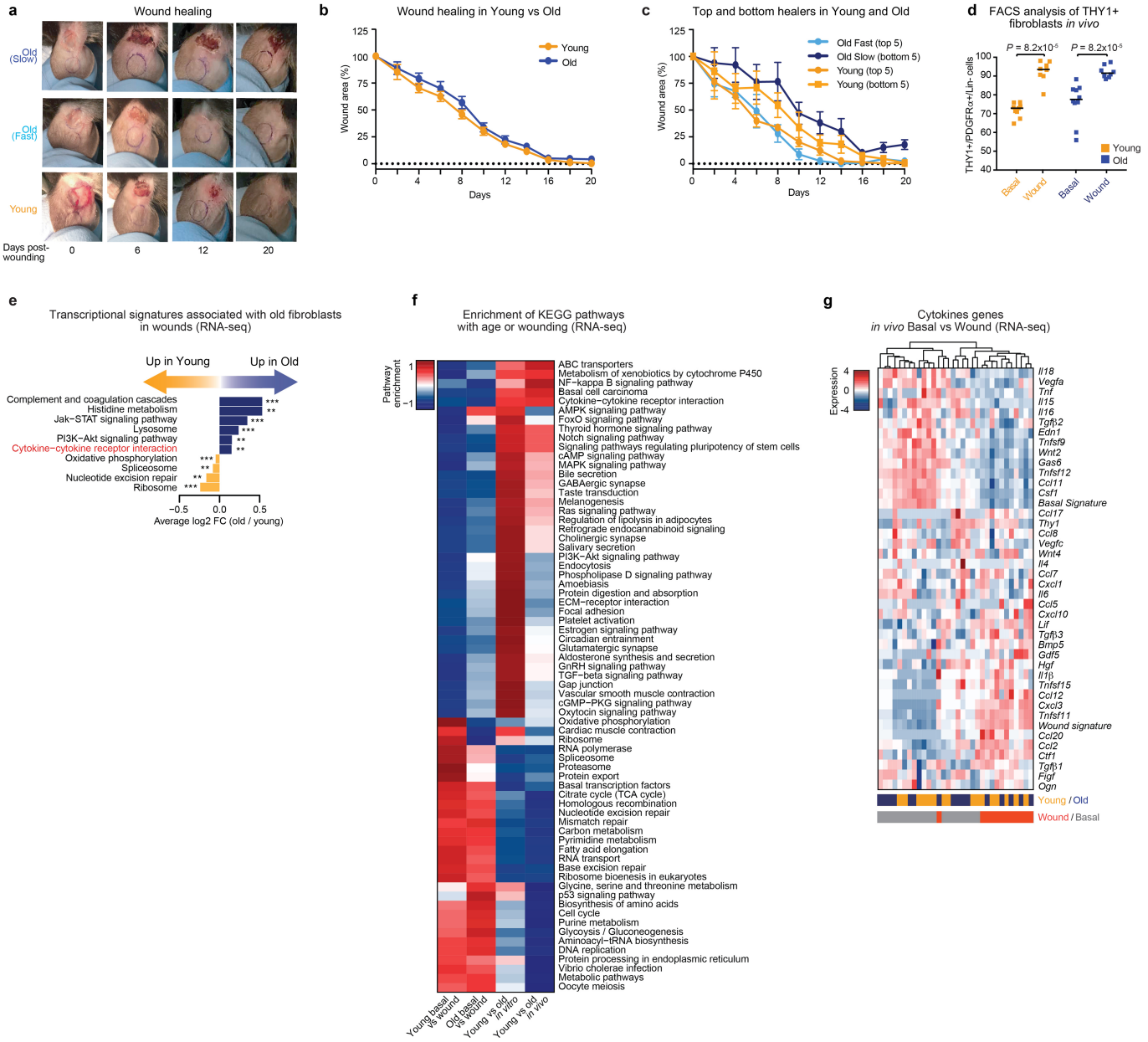

**Extended Data Fig. 5** | See next page for caption.

**Extended Data Fig. 5 | Old fibroblast cultures are enriched for activated (THY1$^+$PDGFRα$^+$) fibroblasts, which are intrinsically poor at reprogramming but facilitate reprogramming extrinsically via secretion of cytokines.** **a**, Left, representative immunofluorescence images of young and old fibroblasts at passage 3 stained for αSMA (which is encoded by the *Acta2* gene) . Right, quantification of the percentage of αSMA$^+$ cells in young (3 months, $n = 8$) and old (29 months, $n = 8$) fibroblasts at passage 3 (1 experiment). Based on reprogramming efficiency, old cultures are shown as good old, bad old or old. Each dot represents cells from one mouse. Lines depict median. *P* value, two-tailed Wilcoxon rank-sum test. **b**, Left, representative immunofluorescence images of young and old fibroblasts at passage 3 incubated with EdU for 4 h, then stained for αSMA (red), EdU (green) and DAPI (blue). White arrows indicate an EdU-positive activated and a non-activated cell. Right, FACS quantification of the percentage of EdU-positive THY1$^-$PDGFRα$^+$ (THY1$^-$) and THY1$^+$PDGFRα$^+$ (THY1$^+$) cells in young (3 months, $n = 5$; 3 independent experiments) and old (29 months, $n = 5$; 1 experiment) cultures at passage 3. Dots and lines as in **a**. *P* values, two-tailed Wilcoxon rank-sum test. **c**, PAGODA of single-cell RNA-seq from young and old fibroblasts described in Extended Data Fig. 4c. Top heat map, PAGODA clustering of cells. Maroon and blue colours indicate increased and decreased expression of the associated gene sets, respectively. Middle heat map, expression of genes in the GO cellular senescence gene set, for which expression is shown as VST-transformed read counts, scaled row-wise. Bottom heat map, cells that originate from good and bad old cultures. The scale for expression fold changes is indicated on the right. **d**, RT–qPCR of *p16$^{Ink4a}$* expression in cultures of THY1$^-$PDGFRα$^+$ (THY1$^-$) and THY1$^+$PDGFRα$^+$ (THY1$^+$) young and old cells at passages 4–6. Results are shown as fold change in expression over THY1$^-$PDGFRα$^+$ cells. Data are from young (3 months, $n = 5$) and old (29 months, $n = 5$) cultures (6 independent experiments). One young and three old cultures were used in 2–3 independent experiments. In this case, an average of the measurements was determined. Dots and lines as in **a**. *P* values, two-tailed Wilcoxon signed-rank test. **e**, Percentage of SA-β-galactosidase-positive cells in young (3 months, $n = 11$) and old (28–29 months, $n = 22$) fibroblast cultures at passage 3 (3 independent experiments). $\log_2$-transformed fold change in SA-β-galactosidase-positive cells over median of young fibroblasts. Line indicates median. *P* values, two-tailed Wilcoxon ranked sum test. **f**, **g**, RT–qPCR of old THY1$^-$PDGFRα$^+$ (THY1$^-$) and THY1$^+$PDGFRα$^+$ (THY1$^+$) cells at passage 4–6 (3 independent experiments) untreated (**f**) ($n = 6$, 3 independent experiments) or treated with the indicated shRNA constructs for 72h (**g**) ($n = 5$, 4 independent experiments). Box-and-whisker plot of fold change in expression over THY1$^-$PDGFRα$^+$ populations originating from the same culture (**f**) or over shLuciferase (shLuc) treated cells (**g**). Box plots depict the median and interquartile range, with whiskers indicating minimum and maximum values. $^*P < 0.06$, one-tailed Wilcoxon signed-rank test with Benjamini–Hochberg correction. **h**, RT–qPCR on young (3 months, $n = 5$) fibroblasts after overexpression of *Ebf2* for 48 h (2 independent experiments). Box-and-whisker plot of fold change in expression over cells treated with empty vector. Box plots as in **f**. $^*P < 0.06$, one-tailed Wilcoxon signed-rank test with Benjamini–Hochberg correction. **i**, Heat map of significantly differentially expressed genes (determined by DESeq2) between freshly FACS-sorted THY1$^-$PDGFRα$^+$Lin$^-$ (THY1$^-$) and THY1$^+$PDGFRα$^+$Lin$^-$ (THY1$^+$) cells described in Fig. 3b and enriched KEGG pathways. Expression is shown as VST-transformed read counts, scaled row-wise. The scale for expression fold changes is indicated on the left. All depicted KEGG pathways were significantly enriched (one-sided Fisher's exact test with Benjamini–Hochberg correction, FDR-adjusted $P < 0.05$). For a complete list of KEGG terms, see Supplementary Table 4e. **j**, Pathway enrichment analysis of KEGG pathways associated with ageing in dataset described in Fig. 3b. For a complete list of KEGG terms, see Supplementary Table 4g. $^{**}P = 0.01$, $^{***}P = 0.001$; two-sided nominal *P* value with Benjamini–Hochberg correction. **k**, Top, ranked fold change (old/young) in levels of the indicated cytokines in plasma (see Extended Data Fig. 1a). Bottom, ranked fold change (old/young) in expression for the indicated cytokines in freshly FACS-sorted THY1$^-$PDGFRα$^+$Lin$^-$ (THY1$^-$) and THY1$^+$PDGFRα$^+$Lin$^-$ (THY1$^+$) cells from young and old ears. See 'Cytokine profiling analysis on plasma and conditioned medium using Luminex multi-analyte' for calculation of ranked fold changes. Gene expression related to wounded fibroblasts is from datasets described in Extended Data Fig. 7e–g. **l**–**n**, Correlation between the proportion of THY1$^+$PDGFRα$^+$ (THY1$^+$) fibroblasts in old cultures (29 months, $n = 23$) (**l**), young (3 months, $n = 21$) and bad old (29 months, $n = 6$) cultures (**m**), and young (3 months, $n = 21$) and good old (29 months, $n = 6$) cultures (**n**), and the reprogramming efficiency of the culture (3 independent experiments). Dots as in **a**. *P* values, two-sided algorithm AS 89 in R. The *y* axis denotes the fold change in the proportion of THY1$^+$PDGFRα$^+$ fibroblasts relative to the median of young mice, and *x* axis denotes the fold change in reprogramming efficiency of the culture relative to the median of young mice. **o**, **p**, Correlation between the proliferation rate (**o**) or the percentage of SA-β-galactosidase-positive cells (**p**) of a given fibroblast culture and reprogramming efficiency of the culture. Proliferation rate was determined by calculating the growth slope of young (3 months, $n = 15$), middle-aged (12 months, $n = 10$) and old (28–29 months, $n = 27$) ear fibroblast cultures at passage 3 (4 independent experiments). Senescence was assessed by SA-β-galactosidase staining of young (3 months, $n = 11$), middle-aged (12 months, $n = 11$) and old (28–29 months, $n = 22$) ear fibroblast cultures at passage 3 (3 independent experiments). Dots as in **a**. *P* values, two-sided algorithm AS 89 in R. The *y* axis denotes the fold change in the proliferation rate or percentage of SA-β-galactosidase-positive cells relative to the median of young mice, and *x* axis denotes the fold change in reprogramming efficiency of the culture relative to the median of young mice. **q**, Reprogramming efficiency of FACS-sorted young (3 months, $n = 8$) and old (29 months, $n = 7$) THY1$^-$PDGFRα$^+$ (THY1$^-$) and THY1$^+$PDGFRα$^+$ (THY1$^+$) fibroblasts at passages 4–6, assessed using AP staining (3 independent experiments). $\log_2$-transformed fold change in reprogramming efficiency of the cells relative to the median of young THY1$^-$PDGFRα$^+$ fibroblasts. One old culture was used in two independent experiments. In this case, an average of the measurements was determined. Dots and lines as in **a**. *P* values, two-tailed Wilcoxon signed-rank test. **r**, Reprogramming efficiency of FACS-sorted young (3 months, $n = 5$) THY1$^-$PDGFRα$^+$ (THY1$^-$) and THY1$^+$PDGFRα$^+$ (THY1$^+$) fibroblasts at passages 4–6, assessed as in **q** (3 independent experiments). Reprogramming was induced using a non-lentiviral piggyBac transposon system. Results are shown as number of AP$^+$ colonies. Dots and lines as in **a**. *P* values, one-tailed Wilcoxon rank-sum test. **s**, **t**, Cytokine profiles of conditioned medium collected from cultures of old (**s**) (29 months, $n = 6$, 3 independent experiments) and young (**t**) (3 months, $n = 6$, 2 independent experiments) THY1$^-$PDGFRα$^+$ (THY1$^-$) and THY1$^+$PDGFRα$^+$ (THY1$^+$) fibroblasts at passages 4–6 . Comparisons were made between THY1$^-$PDGFRα$^+$ and THY1$^+$PDGFRα$^+$ from the same original culture. Based on cytokines that are significantly different in conditioned medium from fibroblasts (Fig. 1b). Box-and-whisker plot of $\log_2$-transformed fold change in mean fluorescence intensity (MFI) over THY1$^-$PDGFRα$^+$ fibroblasts. Box plots as in **f**. $^*P < 0.05$, one-tailed Wilcoxon rank-sum test with Benjamini–Hochberg correction. Exact *P* values are in Supplementary Table 4h. **u**, Reprogramming efficiency, assessed as in **q**, of FACS-sorted young (3 months, $n = 6$) THY1$^-$PDGFRα$^+$ (THY1$^-$) and THY1$^+$PDGFRα$^+$ (THY1$^+$) fibroblasts at passages 4–6 treated with fresh conditioned medium daily starting from day 1 after infection (3 independent experiments). Conditioned medium was collected daily from the THY1$^-$PDGFRα$^+$ or THY1$^+$PDGFRα$^+$ fibroblasts from the same original culture. $\log_2$-transformed fold change in reprogramming efficiency relative to the reprogramming efficiency of THY1$^-$PDGFRα$^+$ fibroblasts treated with conditioned medium from THY1$^-$PDGFRα$^+$ fibroblasts. One young culture was used in two independent experiments. In this case, an average of the measurements was determined. Dots and lines as in **a**. *P* values, two-tailed Wilcoxon signed-rank test. For individual experiments in **b**, **d**–**h**, **l**.**u** and exact *P* values, see Supplementary Table 7.



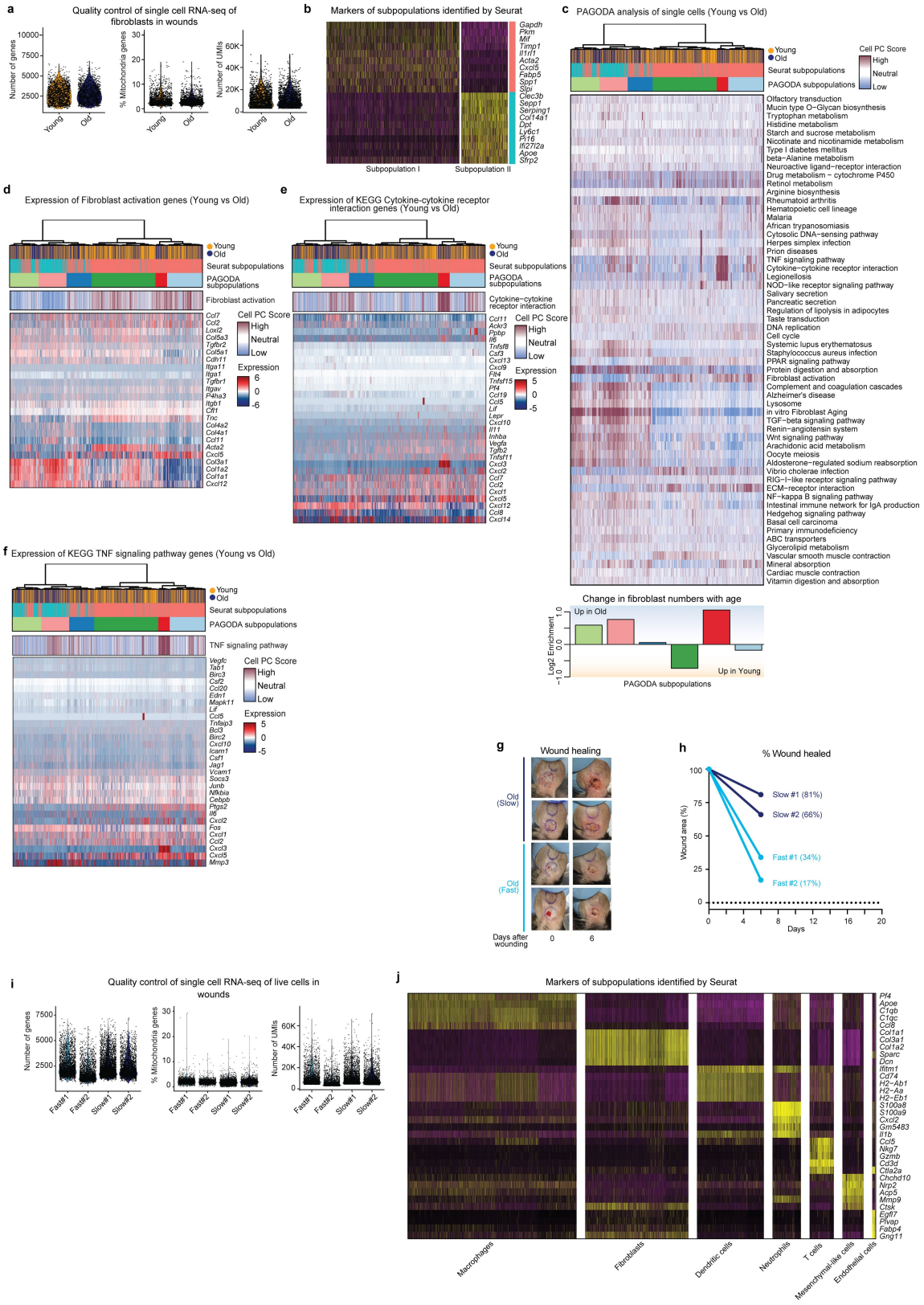

**Extended Data Fig. 6** | See next page for caption.

**Extended Data Fig. 6 | Old fibroblasts secrete cytokines, including IL-6 and TNF, that induce inflammatory signalling pathways and modulate reprogramming efficiency. a**, Reprogramming efficiency, assessed using AP staining, of FACS-sorted young (3 months, $n = 8$) fibroblast cultures at passage 3 that were treated with conditioned medium from day 1 after infection (4 independent experiments). Conditioned medium was collected from young, good old or bad old fibroblast cultures. Results are shown as fold change in reprogramming efficiency relative to young fibroblasts treated with young conditioned medium. Each dot represents cells from one mouse. Lines depict median. *P* values, two-tailed Wilcoxon signed-rank test with Benjamini–Hochberg correction. **b**, Reprogramming efficiency, assessed as in **a**, of bad old (left, $n = 7$) or good old (right, $n = 6$) fibroblast cultures at passage 3 treated with conditioned medium from day 1 after infection (5 independent experiments). Conditioned medium was collected from good or bad old fibroblast cultures. Results are shown as fold change in reprogramming efficiency over old fibroblasts treated with bad old conditioned medium. Dots and lines as in **a**. *P* values, two-tailed Wilcoxon signed-rank test. **c**, Different representation of the data from Fig. 3e; each diamond represents the fold difference in reprogramming efficiency between a unique pair of good and bad old cultures ($n = 8$ pairs of good and bad old cultures, 5 independent experiments). Dots and lines as in **a**. *P* values, two-tailed Wilcoxon signed-rank test with Benjamini–Hochberg correction. **d**, Western blot analysis of young fibroblasts at passage 3 treated with the indicated cytokines at the concentration of 10 ng ml$^{-1}$ for 30 min. Representative of 3 independent experiments. **e**, Reprogramming efficiency, assessed as in **a**, of young fibroblast cultures (3 months, $n = 10$) at passage 3 (3 independent experiments). Cells were treated with the indicated cytokines from day 1 after infection at the concentration of 10 ng ml$^{-1}$. Results are shown as $\log_2$-transformed fold change in reprogramming efficiency over untreated cells. Dots and lines as in **a**. *P* values, two-tailed Wilcoxon signed-rank test with Benjamini–Hochberg correction. **f**, Reprogramming efficiency, assessed by SSEA1 staining, of young fibroblast cultures (3 months, $n = 4$) at passage 3 treated with the indicated cytokines from day 1 after infection at the concentration of 10 ng ml$^{-1}$ (2 independent experiments). Results are shown as $\log_2$-transformed fold change in reprogramming efficiency over untreated cells. Dots and lines as in **a**. *P* values, one-tailed Wilcoxon signed-rank test with Benjamini–Hochberg correction. **g**, Reprogramming efficiency, assessed as in **a**, of old fibroblast cultures (29 months, $n = 7$) at passage 3 treated with the indicated cytokines from day 1 after infection at the concentration of 10 ng ml$^{-1}$ (3 independent experiments). Results are shown as $\log_2$-transformed fold change in the reprogramming efficiency over untreated cells. Note that 1 old culture was used in 2 independent experiments. In this case, an average of the resultant measurements was determined. Dots and lines as in **a**. *P* values, two-tailed Wilcoxon signed-rank test with Benjamini–Hochberg correction. **h**, Reprogramming efficiency, assessed as in **f**, of old fibroblast cultures (29 months, $n = 3$) at passage 3 treated with the indicated cytokines from day 1 after infection at the concentration of 10 ng ml$^{-1}$ (2 independent experiments). Results are shown as $\log_2$-transformed fold change in the reprogramming efficiency over untreated cells. Dots and lines as in **a**. *P* values, one-tailed Wilcoxon signed-rank test with Benjamini–Hochberg correction. **i**, Western blot analysis using the indicated antibodies of young fibroblasts at passage 3 treated with the indicated cytokines (10 ng ml$^{-1}$) and blocking antibodies (8 μg ml$^{-1}$) for 30 min. Cytokines were pretreated with either IgG or their corresponding blocking antibodies for 1 h before treatment. Representative of 2 independent experiments. **j**, Western blot analysis of old fibroblasts at passage 3 treated with the indicated cytokines at a concentration of 10 ng ml$^{-1}$ for 30 min. **k**, Reprogramming efficiency, assessed as in **a**, of young fibroblast cultures (3 months, $n = 4$) at passage 3 treated with the indicated conditions from day 1 after infection (2 independent experiments). Cytokines (10 ng ml$^{-1}$) were pretreated with either IgG or their corresponding blocking antibody (8 μg ml$^{-1}$) for 1 h before treatment. Results are shown as $\log_2$-transformed fold change in reprogramming efficiency over untreated cells. Dots and lines as in **a**. *P* values, one-tailed Wilcoxon signed-rank test with Benjamini–Hochberg correction. **l**, **m**, Reprogramming efficiency, assessed as in **a**, of young fibroblast cultures (3 months, $n = 6$) at passage 3 treated with the indicated conditions from day 1 after infection (3 independent experiments). Conditioned medium was pretreated for 1 h with the indicated blocking antibody before administration. Results are shown as $\log_2$-transformed fold change relative to conditioned medium treated with IgG. Dots and lines as in **a**. *P* values, one-tailed Wilcoxon signed-rank test with Benjamini–Hochberg correction. **n**, Different representation of the data from Fig. 3f; each diamond represents the fold difference in reprogramming efficiency between a unique pair of good and bad old cultures ($n = 6$ pairs of good and bad old cultures, 4 independent experiments). Line marks median. *P* values, two-tailed Wilcoxon signed-rank test with Benjamini–Hochberg correction. **o**, Heat map showing the Spearman rank correlation coefficients between the levels of individual cytokines (top row), the ratio of the levels of TNF to other cytokines (middle row), the ratio of the levels of IL-6 to other cytokines (bottom row), and reprogramming efficiency in young (3 months, $n = 19$) and old (29 months, $n = 18$) cells (2 independent experiments). $^{*}P < 0.05$, $^{**}P < 0.01$; two-sided algorithm AS 89 in R; TNF:IL-6, $P = 0.040$; IL-6:IFNγ, $P = 0.010$; IL-6:IL-18, $P = 0.040$; IL-6:TNF, $P = 0.040$; IL-6:IL-10, $P = 0.040$; IL-6:CSF2, $P = 0.005$. The remaining *P* values can be found in Supplementary Table 7. **p**, **q**, The levels of IL-6 **(p)** or TNF **(q)** are not correlated with reprogramming efficiency. The *y* axis denotes the fold change in the levels of the indicated cytokine relative to the median of young mice and the *x* axis denotes the fold change in reprogramming efficiency of the culture relative to the median of young mice. Data are from young (3 months, $n = 19$) and old (29 months, $n = 18$) mice (2 independent experiments). *P* values, two-sided algorithm AS 89 in R. For individual experiments in **a**–**c**, **e**–**h**, **k**–**q**, see Supplementary Table 7.



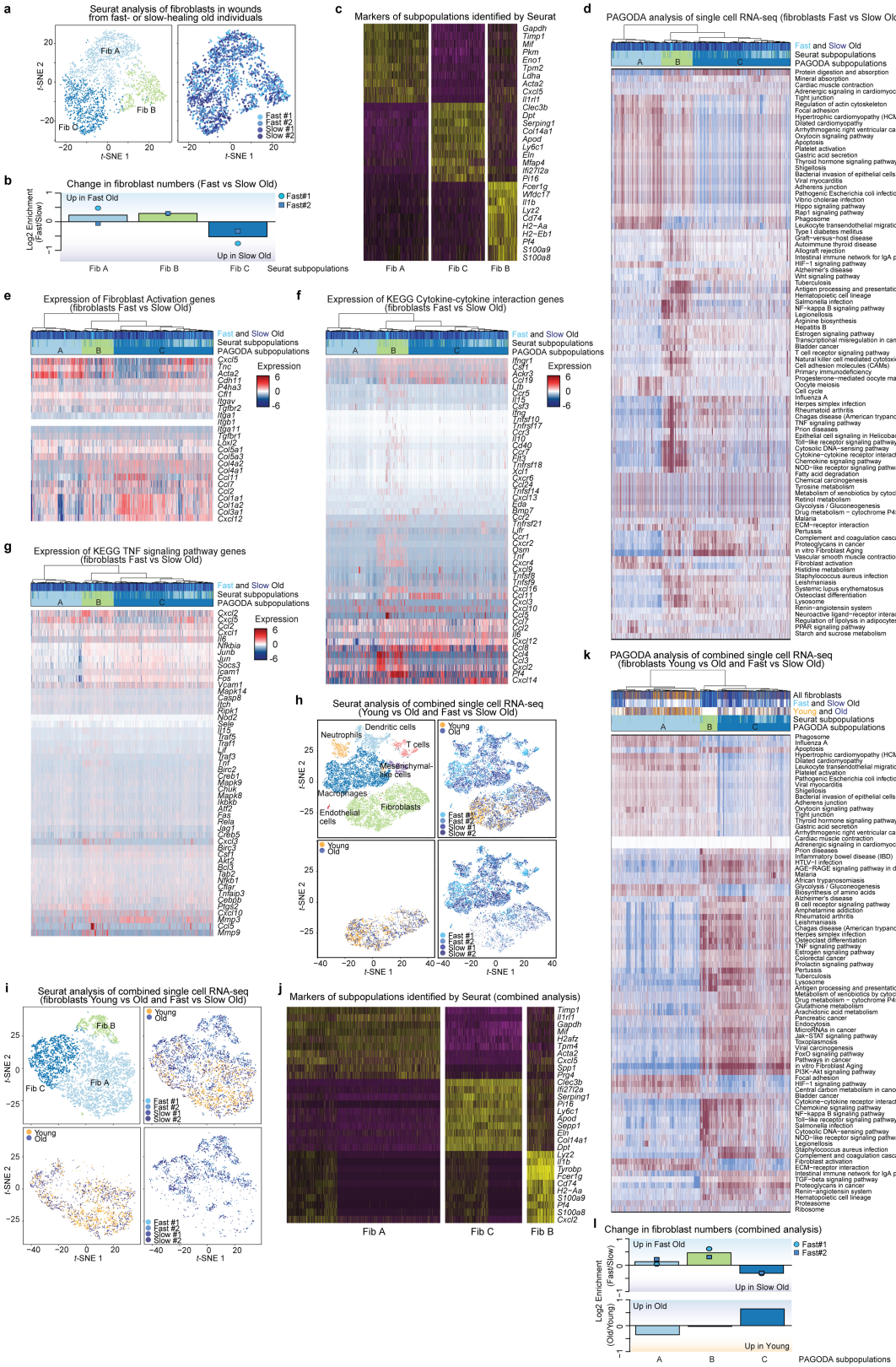

**Extended Data Fig. 7** | See next page for caption.

**Extended Data Fig. 7 | Ageing is associated with an increased variability in wound healing between old mice, and old fibroblasts in wounds are distinct from primary fibroblasts derived from healthy ear skin. a**, Example images of ear wounds of young mice, fast-healing old mice (fast old) and slow-healing old mice (slow old) at the indicated time points (2 independent experiments). Ink circles depict initial size of wounds. **b**, Ear wound healing curve from young (3–4 months, $n = 26$) and old (24–26 months, $n = 28$) mice (2 independent experiments). Full thickness wounds were induced on the dorsal side of both ears (see 'Wounding and wound healing experiments' for details) and the size of the wounds was assessed by imaging ear wounds every second day for 20 days. For each mouse, the average of both ear wounds was calculated. Graph depicts the average percentage of wound area remaining at the indicated time points. Data are mean ± s.e.m. **c**, Ear wound healing curves of the five fastest and the five slowest healing young and old mice. Graph depicts the average average of wound area remaining at the indicated time points. Data are mean ± s.e.m. **d**, FACS analysis as described in Fig. 3b to assess the percentage of $THY1^+PDGFR\alpha^+Lin^-$ ($THY1^+$) cells in ears of young and old mice during basal conditions and at 7 days after induction of wounds. Results are shown as a percentage of $THY1^+PDGFR\alpha^+Lin^-$ cells over $PDGFR\alpha^+Lin^-$ cells. Data shown are from young basal (3–4 months, $n = 9$ replicates, each with 2–3 mice), young wounded (3–4 months, $n = 8$ replicates, each with 2–3 mice), old basal (24–26 months, $n = 10$ replicates, each with 2–3 mice) and old wounded (24–26 months, $n = 8$ replicates, each with 2–3 mice) (3 independent experiments). Each dot represents a replicate with cells pooled from 2–3 mice. Line depicts median percentage. *P* values, two-tailed Wilcoxon rank-sum test. Note that the percentage of $THY1^+PDGFR\alpha^+Lin^-$ in young and old basal conditions is also presented in Fig. 3b. **e**, Pathway enrichment analysis based on population RNA-seq of young wounded (3–4 months, $n = 6$ replicates, each with 2–3 mice) and old wounded (24–26 months, $n = 6$ replicates, each with 2–3 mice) $THY1^-PDGFR\alpha^+Lin^-$ and $THY1^+PDGFR\alpha^+Lin^-$ cells in vivo (1 experiment). The graph shows a subset of KEGG pathways that were found to be significantly enriched (FDR-adjusted $P < 0.05$). For a complete list of differentially expressed genes and pathways, with corresponding specific *P* values, see Supplementary Table 5c, d. $^{**}P < 0.01$, $^{***}P < 0.001$; two-sided nominal *P* value with Benjamini–Hochberg correction. **f**, Comparison between the transcriptomic changes that occur in fibroblasts with age in vitro (as described in Fig. 2b) and in vivo (as described in Fig. 3b), as well as changes that occur upon wounding in young and old ears (as described in **d**). The heat map depicts the enrichment of the KEGG pathways that are present in at least two of the conditions described. For the complete list of differentially expressed genes and significant KEGG terms with specific *P* values, see Supplementary Tables 2b, c, 4f, g, 5a, b. The scale for enrichment is indicated on the left. **g**, Heat map of expression of a subset of cytokine genes from population RNA-seq of fibroblasts from young and old ears during basal and wounded conditions. Expression is shown as VST-transformed read counts, scaled row-wise. The scale for expression fold changes is indicated on the left. Basal and wound signatures refer to the average expression of the genes that are significantly downregulated or upregulated with wounding, respectively, in this dataset.



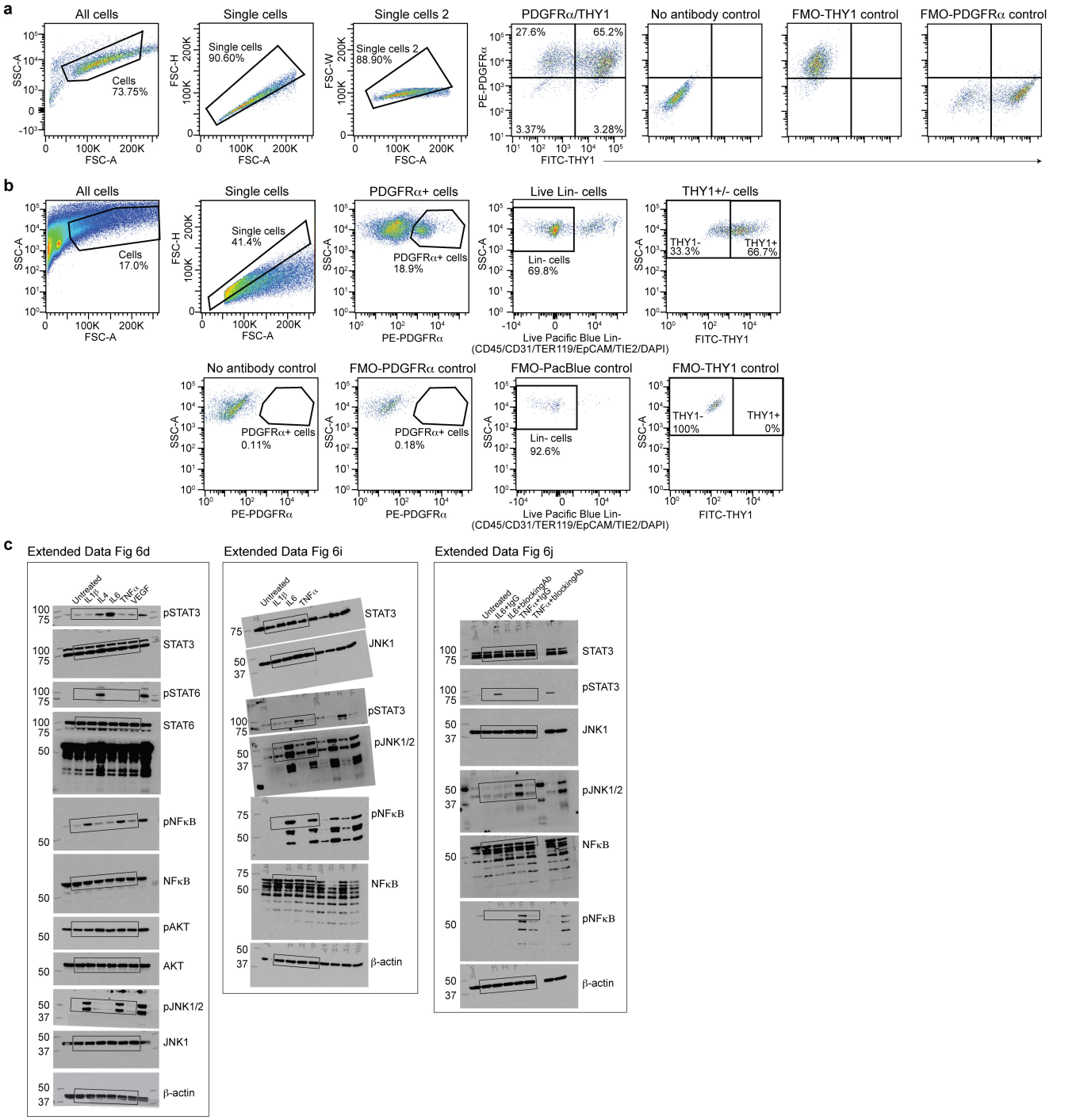

**Extended Data Fig. 8** | See next page for caption.

**Extended Data Fig. 8 | Single-cell RNA-seq analysis of fibroblasts in wounds from young and old mice and single-cell RNA-seq analysis of entire wounds from old slow- and fast-healing mice at day 7. a**, Quality control for 10x Genomics single-cell RNA-seq data of freshly isolated PDGFRα$^+$Lin$^-$ (CD45$^-$CD31$^-$EpCAM$^-$TER119$^-$TIE2$^-$) cells from wounds of young mice (3–4 months, cells pooled from $n = 10$ mice) or old mice (24–26 months, cells pooled from $n = 10$ mice), 7 days after induction of wounds. Number of unique genes (left), percentage of mitochondrial genes (middle) and number of unique molecular identifiers (UMIs, right) for each cell are shown, separated by age group. Each dot represents a single cell. **b**, Seurat analysis of all live high-quality PDGFRα$^+$Lin$^-$ cells described in **a** (3,036 cells in total) identified two main clusters of cells. Heat map depicts the expression of the top 10 marker genes for each significant cell cluster identified by Seurat, which are defined as the genes that are most specific to each population. The cell subpopulation identity assigned to each cluster is indicated below each column. **c**, PAGODA of the single-cell RNA-seq dataset described in **a**. PAGODA was performed using all KEGG pathways, and the in vitro fibroblast ageing and the fibroblast activation signatures (see Supplementary Table 2b, f). Top, heat map of single cells from wounds from young and old mice and cell clusters identified by Seurat and PAGODA analyses. Bottom, heat map of the separation of cells based on their principal component scores for the significantly overdispersed gene sets. Top heat map, PAGODA clustering of cells. Maroon and blue colours indicate increased and decreased expression of the associated gene sets, respectively. Bottom, $\log_2$-transformed fold change in the subpopulations between young and old wounds at day 7. **d**, PAGODA as described in **c**. Middle, heat map of the separation of cells based on their principal component scores for the fibroblast activation signature. Bottom, heat map of the expression of the genes that are part of the fibroblast activation signature (see Supplementary Table 2f); expression is shown as log-transformed and normalized gene expression values as calculated by Seurat and scaled row-wise. The scale for expression fold changes is indicated on the right. **e**, PAGODA as described in **c**. Middle, heat map of the separation of cells based on their principal component scores for the KEGG cytokine–cytokine receptor interaction gene set. Bottom, heat map of the expression of the genes that are part of the KEGG cytokine–cytokine receptor interaction; expression is shown as log-transformed and normalized gene expression values as calculated by Seurat and scaled row-wise. The scale for expression fold changes is indicated on the right. **f**, PAGODA as described in **c**. Middle, heat map of the separation of cells based on their principal component scores for the KEGG TNF signalling pathway gene set. Bottom, heat map of the expression of the genes that are part of the KEGG TNF signalling pathway; expression is shown as log-transformed and normalized gene expression values as calculated by Seurat and scaled row-wise. The scale for expression fold changes is indicated on the right. **g**, Representative images of the ears of the two slow-healing and two fast-healing old mice used for single-cell RNA-seq at day 7 after wounding (1 experiment). **h**, Ear wound healing curves of the ears of the two slow-healing and two fast-healing old mice used for single-cell RNA-seq. The percentage of the wound area that was not healed at the day 6 after induction of the wounds is indicated in parentheses. **i**, Quality control for 10x Genomics single-cell RNA-seq data of freshly isolated live cells from the ear wounds of slow-healing ($n = 2$) and fast-healing old mice ($n = 2$), 7 days after induction of wounds. Number of genes (left), percentage of mitochondrial genes (middle) and number of unique molecular identifier (right) for each cell are shown, separated by mouse. Each dot represents a single cell. **j**, Seurat analysis of all live high-quality cells described in **i** (10,797 cells in total) identified seven main clusters of cells. Heat map depicts the expression of the top 10 marker genes for each significant cell cluster identified by Seurat, which are defined as the genes that are most specific to each population. The cell subpopulation identity assigned to each cluster is indicated below each column.




**Extended Data Fig. 9** | See next page for caption.

**Extended Data Fig. 9 | Seurat and PAGODA single-cell RNA-seq analyses of fibroblasts identify distinct fibroblast subpopulations associated with fast- or slow-healing trajectories. a**–**g**, Analysis of cells identified as fibroblasts from the single-cell RNA-seq dataset described in Fig. 4c. **a**, *t*-SNE clustering of cells identified as fibroblasts (2,678 cells in total) coloured by significant clusters identified using a *k*-nearest neighbour (KNN) graph-based algorithm as implemented by Seurat, or by mouse. **b**, $\log_2$-transformed fold change in the number of cells in each of the three subpopulations identified by Seurat between fast-healing old wounds and slow-healing old wounds. **c**, Seurat analysis of fibroblasts (2,678 cells in total) identified three main clusters. Heat map depicts the expression of the top 10 marker genes for each significant cell cluster identified by Seurat, which are defined as the genes that are most specific to each population. The identity of each cell subpopulation assigned to each cluster is indicated below each column. **d**, PAGODA of fibroblasts. PAGODA was performed using raw expression counts and all KEGG pathways, and the in vitro fibroblast ageing and the fibroblast activation signatures (see Supplementary Table 2b, f). Top, heat map of single cells from wounds of old mice and cell clusters identified by Seurat and PAGODA analyses. Bottom, heat map of the separation of cells based on their principal component scores for the significantly overdispersed gene sets. Top heat map, PAGODA clustering of cells. Maroon and blue colours indicate increased and decreased expression of the associated gene sets, respectively. **e**, PAGODA as described in **d**. Bottom, heat map of the expression of the genes that are part of the fibroblast activation signature (see Supplementary Table 2f); expression is shown as log-transformed and normalized gene expression values as calculated by Seurat and scaled row-wise. The scale for expression fold changes is indicated on the right. **f**, Expression of the genes that are part of the KEGG cytokine–cytokine receptor interaction gene set as in **e**. **g**, Expression of the genes that are part of the KEGG TNF signalling pathway as in **e**. **h**–**l**, Analysis of the combined single-cell RNA-seq datasets described in Fig. 4b, c. **h**, Seurat analysis of combined datasets clusters fibroblasts from both datasets together. *t*-SNE clustering of all live, high-quality cells from both datasets (13,833 cells in total) coloured by significant clusters identified using a KNN graph-based algorithm as implemented by Seurat, or by mouse. **i**, *t*-SNE clustering of combined fibroblasts from the datasets described in Fig. 4b (PDGFRα$^+$Lin$^-$) and Fig. 4c. Combined fibroblasts (5,716 cells in total) are coloured by significant clusters identified using a KNN graph-based algorithm as implemented by Seurat, or by mouse. **j**, Seurat analysis of combined fibroblasts (5,716 cells in total) identified three main subpopulations. Heat map depicts the expression of the top 10 marker genes for each significant subpopulation identified by Seurat, which are defined as the genes that are most specific to each population. The cell subpopulation identity assigned to each cluster is indicated below each column. **k**, PAGODA of combined fibroblasts. PAGODA was performed using Seurat normalized counts and all KEGG pathways, the in vitro fibroblast ageing, the fibroblast activation signatures (see Supplementary Table 2b, f). Top, heat map of single fibroblasts from wounds of young and old mice or wounds from old fast- or slow-healing mice, and cell clusters identified by Seurat and PAGODA analyses. Bottom, heat map of separation of cells based on their principal component scores for a subset of the top significantly overdispersed gene sets. Top heat map, PAGODA clustering of cells. Maroon and blue colours indicate increased and decreased expression of the associated gene sets, respectively. Note that fibroblast subpopulation B did not contain cells from old/young in the combined analysis. This is probably owing to the fact that this subpopulation of fibroblast has some markers of the haematopoietic lineage, and is probably depleted in the PFGDRα$^+$Lin$^-$ FACS-sorting scheme used to isolate fibroblasts from the wounds of young and old mice. **l**, $\log_2$-transformed fold change in each of the three combined fibroblast subpopulations identified by PAGODA between wounds of old fast- and slow-healing mice, or between wounds from young and old mice, at day 7.




**Extended Data Fig. 10 | FACS schematic for in vitro and in vivo fibroblast analysis and sorting and full western blot membranes. a**, FACS schematic for analysis and sorting of THY1$^-$PDGFRα$^+$ and THY1$^+$PDGFRα$^+$ cells in young and old cultures at passage 3. Gates shown on each plot are indicated above the plot. Marker and fluorophore are shown on each axis. FMO, fluorescence minus one. **b**, FACS schematic for analysis and sorting of live THY1$^-$PDGFRα$^+$Lin$^-$ and THY1$^+$PDGFRα$^+$Lin$^-$ cells from young and old fresh tissues (ears), used for population RNA-seq and single-cell RNA-seq analyses. Gates shown on each plot are indicated above the plot. Marker and fluorophore are shown on each axis. **c**, Full western blot membranes from Extended Data Fig. 6d, i, j. Boxes indicate the cropped area.

nature research

Corresponding author(s): Anne Brunet

Last updated by author(s): Aug 16, 2019

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| n/a | Confirmed |
|---|---|
| ☐ | ☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☐ | ☒ A description of all covariates tested |
| ☐ | ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☒ | ☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☒ | ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☐ | ☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |

*Our web collection on statistics for biologists contains articles on many of the points above.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | Imaging: AxioVision 4.7.2, NIS Elements AR software (v4.30.02).<br>FACS: BD FACSDiva software (v8.0.1)<br>qRT-PCR: Bio-Rad CFX manager (v3.1) |
| Data analysis | Data analysis was performed using R version 3.2.1, 3.3.1, or 3.5.0. Key packages used were: DESeq2 (v1.20.1, v1.6.3), EdgeR (v3.10.2), DiffBind (v.1.12.3), scde (v0.99.1, v2.8.0), cellrangerRKit (v1.1.0), ggplot2 (version 3.0.0), FlowJo (version 10.2), ImageJ (v1.47), Seurat (v2.3.4), trim-galore software (v0.2.1), bowtie v0.12.7, FIXSEQ, MACS (v2.08), TopHat (v2.0.8b), and HTSeq (v0.6.1), Enricher (http://amp.pharm.mssm.edu/Enrichr/), QIAGEN's Ingenuity Pathway analysis (IPA QIAGEN Redwood City). |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All sequencing data have been deposited in NCBI BioProject database under the accession code PRJNA316110. Figures 2, 3, and 4 of this study are all associated with raw data, which can be found under this accession number. For Figures 2-3, raw FASTQ files for RNA-seq and ChIP-seq are provided. For Figure 4, barcoded BAM files for 10x single cell RNA-sequencing are provided.

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Sample size | Except for the wound healing experiment, no statistical analysis was performed to pre-determine sample size (most of the sample size chosen are standard for the field). This is clearly indicated in Experimental Procedures (Statistical analysis section). In cases where samples from independent experiments were combined, we have clearly indicated this, and the non-combined data are provided in a Supplementary Table 7. For the wound healing experiment, a power analysis was performed based on an initial experiment to determine the sample size required to detect a difference in variability with a 95% confidence interval. |
| Data exclusions | Samples were excluded due to pre-established exclusion criteria (QC features), if associated phenotypic data was not reproducible, or due to batch-effects that could not be corrected for.<br><br>Specifically, two plasma samples from old mice were discarded, as the coefficient of variation (CV) was > 20% for most of the cytokines measured between the two technical replicates for these two plasma samples. The following RNA-seq samples were excluded from further analyses: (1) 2 old and 3 middle-aged RNA-seq libraries that lacked corresponding young samples, and could therefore not be corrected for batch; (2) RNA-seq libraries from 1 good old and 1 bad old fibroblast cultures whose reprogramming efficiency could not be confirmed over several independent experiments; (3) RNA-seq libraries from 2 iPSC lines (out of 13 total) failed at the QC stage because they exhibited large differences (for example in number of reads mapped) from the rest of the samples (Supplementary Table 1f). For the wound healing experiment, 2 old mice were excluded, as they died before day of wound closure could be determined. This information is stated in the corresponding sections in Experimental Procedures. |
| Replication | All attempts at replication were successful. Independent replication was done for: Fig. 1b, c, Fig. 3a, b, c, d, e, f, g, Fig. 4a, Extended Data Fig. 1a, b, e, l, m, n, Extended Data Fig. 5b, d, e, f, g, h, l, m, n, o, p, q, r, s, t, u, Extended Data Fig. 6a, b, c, d, e, f, g, h, i, k, l, m, n, o, p, q, Extended Data Fig. 7a, b, c, d, Extended Data Fig. 10a, b, c. All of the independent experiments are presented in a supplementary table (Supplementary Table 7, except images such as western blots) and this is indicated in figure legends. Independent replication was not done for: Fig. 2b, c, d, e, f, Fig. 4b, c, d, Extended Data Fig. 1c, f, g, j, k, o, Extended Data Fig. 2, Extended Data Fig. 3, Extended Data Fig. 4, Extended Data Fig. 5a, c, i, j, Extended Data Fig. 7e, f, g, Extended Data Fig. 8, Extended Data Fig. 9. The “omics” datasets comprise independent biological samples, but were not independently replicated (RNA-seq, Single-cell RNA-seq, ChIP-seq and metabolomics datasets). Human fibroblast experiments, experiments with fibroblasts at passage 33, and generation of iPSC lines were not independently replicated. This is stated in the corresponding figure legends. |
| Randomization | For the majority of experiments, young and old mice were processed in an alternating manner rather than in two large groups to minimize group effect. This is indicated in Experimental Procedures (Statistical analysis section). |
| Blinding | For all quantifications that were done with FACS or automated image quantification, no blinding was performed, including Fig. 3a, b, c, Extended Data Fig. 5l, m, n, Extended Data Fig. 7d. The other experiments were blinded, with the exception of Fig. 1c, Extended Data Fig. 1l, m, Extended Data Fig. 5e, Extended Data Fig. 6f, g, h, k. This is indicated in Experimental Procedures (Statistical analysis section). |

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials & experimental systems

| n/a | Involved in the study |
|---|---|
| ☐ | ☒ Antibodies |
| ☐ | ☒ Eukaryotic cell lines |
| ☒ | ☐ Palaeontology |
| ☐ | ☒ Animals and other organisms |
| ☐ | ☒ Human research participants |
| ☒ | ☐ Clinical data |

## Methods

| n/a | Involved in the study |
|---|---|
| ☐ | ☒ ChIP-seq |
| ☐ | ☒ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |

## Antibodies

| | |
|---|---|
| Antibodies used | Antibodies used for FACS analysis: PE-CD140a (BioLegend, #135905, Clone: APA5, Lot#B218257/B244566, [1:100]), FITC-CD90.2 |

(BioLegend, #105305, Clone: 30-H12, Lot#B184407/B224687, [1:200]), Brilliant Violet 421-TER119 (BioLegend, #116234, Clone: TER-119, Lot#B244217, [1:100]), eFluor 450-CD326 (Fisher Scientific, #50-163-76, Clone: G8.8, Lot#4277888, [1:200]), Pacific Blue-CD45 (BioLegend, #103126, Clone: 30-F11, Lot#B242848/B253970, [1:200]), Pacific Blue-CD31 (BioLegend, #102422, Clone: 390, Lot#B182438, [1:100]), Biotin-CD202b (Thermo Fisher Scientific, #13-5987-82, Clone: TEK4, Lot#B231548, [1:200]), Brilliant Violet 421-Streptavidin (BioLegend, #405226, Lot#B240413, [1:200]), APC-B220 (eBioscience, #47-0452-82, Clone: RA3-6B2, Lot#NN, [1:100]), APC-CD3 (BD Pharmingen, #557597, Clone: SP34-2, Lot#NN, [1:100]), APC-Gr-1 (eBioscience, #17-5931-82, Clone: RB6-8C5, Lot#NN, [1:100]), APC-F4/80 (eBioscience, #17-4801-82, Clone: BM8, Lot#NN, [1:100]), APC-Siglec H (BioLegend, #129611, Clone: 551, Lot#NN, [1:100]), APC-CD11c (eBioscience, #17-0114-82, Clone: N418, Lot#NN, [1:100]).

Antibodies used for immunofluorescence: OCT3/4 (Santa Cruz Biotech, #sc9081, Polyclonal, Lot#L2211, [1:200]), SSEA-1 StainAlive DyLight488 (Stemgent, #09-0067, Clone: MC-480, Lot#J16010000000009/2482, [1:100-200]), SOX2 (Santa Cruz Biotech, #sc17320, Polyclonal, Lot#, [1:200]), αSMA (Abcam, #ab7817, Clone: 1A4, Lot#GR119216-7, [1:2000]), and Alexa Fluor 488 αSMA (Abcam, #ab184675, Clone: 1A4, Lot#GR316286-2, [1:200]).

Antibodies used for blocking experiments: IgG (R&D systems, #AB-108-C, Polyclonal, Lot#ES4116081/ES4115041/ES4114041, [8 µg/mL]), IL6 (R&D systems, #AB-406-NA, Polyclonal, Lot#BF0916041/BF0913111, [8 µg/mL]), and TNFα (R&D systems, #AB-410-NA, Polyclonal, Lot#CT0714031/CT0715031/CT0716021, [8 µg/mL]).

Antibodies used for Western blot: phospho-STAT3 (Tyr705) (Cell Signaling Technology, #9145, Clone: D3A7, Lot#26, [1:1000]), STAT3 (Invitrogen, #44-364G, Clone: 44-364G, Lot#0601, [1:2000]), phospho-STAT6 (Tyr641) (Cell Signaling Technology, #9361, Polyclonal, Lot#12, [1:1000]), STAT6 (Cell Signaling Technology, #5397, Clone: D3H4, Lot#1, [1:1000]), phospho-AKT (Ser473) (Cell Signaling Technology, #4060, Clone: D9E, Lot#19, [1:2000]), AKT (Cell Signaling Technology, #4691, Clone: C67E7, Lot#20, [1:1000]), phospho-NFκB (Ser536) (Cell Signaling Technology, #3033, Clone: 93H1, Lot#14, [1:1000]), NFκB (Cell Signaling Technology, #8242, Clone: D14E12, Lot#4, [1:1000]), phospho-JNK1&2 (Thr183 & Tyr185) (Invitrogen, #44-682G, Polyclonal, Lot#RC220615, [1:1000]), JNK1 (Invitrogen, #44-690G, Polyclonal, Lot#RC222625, [1:2000]), and β-actin (Novus Biologicals, #NB600-501, Clone: AC-15, Lot#061M4808, [1:50,000]).

Antibodies used for ChIP experiments: H3K4me3 (Active Motif, #39159, Polyclonal, Lot#NN, 5ug), and H3K27me3 (Active Motif, #39536, Clone: 7B11, Lot#NN, 5ug).

Antibodies used for MACS: APC PSA-NCAM (Miltenyi, #130-093-273, Clone: 2-2B, Lot#NN, [1:8]).

Secondary antibodies: HRP-conjugated goat anti-mouse (Calbiochem, #401215, Lot#D00157542, [1:5000]), HRP-conjugated goat anti-rabbit (Calbiochem, #401393, Lot#D00168510, [1:5000]), donkey anti-mouse AF568 (ThermoFisher, #A10037, Lot#1752099, [1:500]), donkey anti-rabbit AF568 (ThermoFisher, #A10042, Lot#1964370, [1:500]).

NN= Not noted

Validation

Only commercially available antibodies that have been widely cited in the literature were used in this study. Antibody specificity and quality validation were performed by the manufacturers (see manufacturers' webpages for further information).

The following additional validations were performed:
-The CD140a (PDGFRa) and CD90 (THY1) antibodies were validated by confirming the expression/lack of expression of the respective genes by RNA-seq in the FACS-sorted populations.
-The OCT3/4 and SOX2 staining were nuclear as expected.
-Antibodies used for blocking experiments (polyclonal IL6 and TNFα antibodies) were validated by testing whether they block the activity of corresponding cytokine in in vitro assays.
-Antibodies used for western blotting gave a band of the expected molecular weight.
-Phospho-antibodies used for western blotting exhibited increased signal in samples treated with cytokines that are known to induce phosphorylation of the corresponding protein (e.g. IL6 treatment induced phospho-STAT3, TNFa treatment induced phospho-NFκB and phospho-JNK1&2).

## Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

293T (ATCC, #CRL-11268)

Authentication

Cell line was not authenticated in-house, but 293T cells were purchased from ATCC where routine cell authentication is conducted. 293T cells were used only at early passages (< passage 20) for lentiviral production.

Mycoplasma contamination

293T cells were negative for mycoplasma. Mycoplasma testing was conducted at regular intervals (2-3 months).

Commonly misidentified lines
(See ICLAC register)

-

## Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals

All mice used in this study were male C57BL/6 obtained from the NIA Aged Rodent colony with ages ranging from 3-4 months for young adult animals, 12-13 months for middle-aged animals, and 20-30 months for old animals (precise ages are stated for each experiment in Supplementary Data Table 7). Mice were habituated for >1 week at Stanford before use. At Stanford, all mice

| | |
|---|---|
| | were housed in the Comparative Medicine Pavilion, and their care monitored by the Veterinary Service Center at Stanford University under IACUC protocol #8661. |
| Wild animals | No wild animals were used in this study. |
| Field-collected samples | No field collection of samples was conducted in this study. |
| Ethics oversight | At Stanford, all mice were housed in the Comparative Medicine Pavilion, and their care monitored by the Veterinary Service Center at Stanford University under IACUC protocol #8661. |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Human research participants

Policy information about studies involving human research participants

| | |
|---|---|
| Population characteristics | As stated in Experimental Procedures, biopsies were collected from male participants of different ages (ranging from 25-90 years old, Supplementary Table 1g) with four biological grandparents of Ashkenazi Jewish descent, generally healthy without thyroid disease, diabetes, immunodeficiency, ongoing cancer or autoimmune disease, and no history of poor wound healing. |
| Recruitment | All participants were recruited using paper postings or word of mouth. To restrict confounds of genetic background and disease state, participants were required to have four biological grandparents of Ashkenazi Jewish decent, and to be generally healthy without thyroid disease, diabetes, immunodeficiency, ongoing cancer or autoimmune disease, and no history of poor wound healing. Hence, there is a bias in genetic background and disease state. |
| Ethics oversight | Stanford Human Subjects approval and informed consent was obtained prior to all study procedures. |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## ChIP-seq

### Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

| | |
|---|---|
| Data access links<br>*May remain private before publication.* | All raw sequencing reads for ChIP-seq data can be found under BioProject PRJNA316110. |
| Files in database submission | The files are annotated as:<br>Young_Fib_culture4_H3K27me3<br>Young_Fib_culture5_H3K27me3<br>Old_Fib_culture4_H3K27me3<br>Old_Fib_culture5_H3K27me3<br>Old_Fib_culture6_H3K27me3<br><br>Young_Fib_culture4_H3K4me3<br>Young_Fib_culture5_H3K4me3<br>Old_Fib_culture4_H3K4me3<br>Old_Fib_culture5_H3K4me3<br>Old_Fib_culture6_H3K4me3<br><br>Young_Fib_culture4_INPUT<br>Young_Fib_culture5_INPUT<br>Old_Fib_culture4_INPUT<br>Old_Fib_culture5_INPUT<br>Old_Fib_culture6_INPUT |
| Genome browser session<br>(e.g. UCSC) | http://genome.ucsc.edu/cgi-bin/hgTracks?<br>hgS_doOtherUser=submit&hgS_otherUserName=salahm&hgS_otherUserSessionName=Mahmoudi_et_al_2018 |

### Methodology

| | |
|---|---|
| Replicates | ChIP-seq (H3K4me3 and H3K27me3) and input (10% of chromatin used for each ChIP reaction) libraries were generated for 2 young and 3 old independent fibroblast cultures. |
| Sequencing depth | ChIP and input libraries were sequenced on Illumina HiSeq 2000 platform (single-end 50bp reads) Table below indicated the number of reads acquired for each sample and quality-control measures.<br><br>Sample Molecule Reads Read Length FIXSEQ unique reads PCR dulplication rate (FIXSEQ) MACS2 peaks (FDR <1%) Peaks with >5 fold enrichment (MACS2)<br>Old_Fib_culture4_H3K27me3 H3K27me3 43662273 50 27670751 29.28 48741 29084<br>Old_Fib_culture4_H3K4me3 H3K4me3 21847330 50 14813404 17.26 25543 24571 |

Old_Fib_culture4_INPUT INPUT 20848880 50 16879828 1.5 NA NA
Old_Fib_culture5_H3K27me3 H3K27me3 49256681 50 35092468 20.82 100354 68123
Old_Fib_culture5_H3K4me3 H3K4me3 23314898 50 17070330 10.76 24556 23557
Old_Fib_culture5_INPUT INPUT 21550022 50 17799159 1.29 NA NA
Old_Fib_culture6_H3K27me3 H3K27me3 50766024 50 36941549 19.12 79715 52861
Old_Fib_culture6_H3K4me3 H3K4me3 23894584 50 16569260 16.34 24882 24162
Old_Fib_culture6_INPUT INPUT 20205828 50 15692624 7.18 NA NA
Young_Fib_culture4_H3K27me3 H3K27me3 49198534 50 34228562 22.21 100087 66305
Young_Fib_culture4_H3K4me3 H3K4me3 23664469 50 13669641 29.49 23354 22605
Young_Fib_culture4_INPUT INPUT 20248064 50 15203520 7.93 NA NA
Young_Fib_culture5_H3K27me3 H3K27me3 39027396 50 23785464 31.32 25087 29084
Young_Fib_culture5_H3K4me3 H3K4me3 21376578 50 12492647 31.72 21843 21373
Young_Fib_culture5_INPUT INPUT 18897473 50 15341132 1.32 NA NA

Antibodies

Antibodies used for ChIP experiments: H3K4me3 (Active Motif, #39159), H3K27me3 (Active Motif, #39536).

Peak calling parameters

Trimmed reads were mapped to the mm9 genome assembly using bowtie v0.12.7. Duplicate reads were eliminated using the FIXSEQ software with default parameters ChIP-seq peaks were called in all samples using the MACS (v2.08) software with default settings and the "--broad" option. Input datasets were used as baseline.

Data quality

For analysis, Fastq reads were quality-trimmed using the trim-galore software (v0.2.1), with a Phred score threshold of 15, and a minimum remaining read length of 36bp. The number of peaks for each sample at FDR 0.01, and above 5-fold enrichment is indicated in the table above.

Software

The following packages were used for analyses: trim-galore software (v0.2.1), bowtie (v0.12.7), FIXSEQ, and MACS (v2.08).

# Flow Cytometry

## Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

## Methodology

Sample preparation

For FACS quantification and sorting of fibroblast subpopulations from primary fibroblast cultures
We performed FACS analysis and sorting of THY1+/ PDGFRα+ and THY1-/ PDGFRα+ cells from primary fibroblast cultures at passage 3. Cells were dissociated and resuspended in FACS buffer (1% BSA, 1mM EDTA in PBS), and stained for the cell surface markers PE-CD140a (BioLegend, #APA5) and FITC-CD90.2 (BioLegend, #30-H12) for 30 minutes on ice. For EdU incorporation experiments, fibroblast cultures were assessed by FACS using the Click-iT EdU Plus FACS PacBlue Kit (Invitrogen, #C10636) in accordance with manufacturer's instructions. Briefly, fibroblasts were incubated in media containing 5-ethynyl-2'-deoxyuridine (EdU; 10 µM) for 4 hours. Cells were then dissociated and resuspended in FACS buffer (1% BSA in PBS). Cell surface markers were stained with PE-CD140a (BioLegend, #APA5) and FITC-CD90.2 (BioLegend, #30-H12). Cells were then fixed (4% paraformaldehyde, PBS) and permeabilized, followed by click reaction to detect EdU, according to the manufacturer's instructions. FACS analysis was performed using a on a LSR II flow cytometer (BD Biosciences), and FACS sorting was performed on a BD FACS Aria II sorter, using a 100 µm nozzle at 13.1 PSI.

For FACS quantification and sorting of fibroblasts from fresh tissues
We isolated fibroblasts from the ears of young and old mice for FACS for quantification and for transcriptomic analysis at population level. Ears of 2-3 mice were dissected and pooled for each sample, cut into small fragments (~1 mm2), and then digested in Dulbecco's Modified Eagle Medium (DMEM, Invitrogen, #11965-092) supplemented with 0.14 Wunsch units/mL of Liberase DL (Roche, # 5401160001) for 30 minutes at 37°C. The fragments were washed with DMEM supplemented with 20% fetal bovine serum (FBS, Gibco, #16000-044), funneled through a 100 µm nylon mesh (Fisher Scientific, #08-771-19), and washed with fibroblast growth medium (DMEM supplemented with 10% FBS and 1% PSQ). A second filtering was performed using a 40µm nylon mesh (Fisher Scientific, #08-771-1), followed by a washing step with fibroblast growth medium. Finally, cells were washed with FACS buffer (PBS, 1% BSA, 500nM EDTA), and then re-suspended in FACS buffer and ready to be stained for FACS analysis. For in vivo FACS analysis and sorting the following antibodies were used: CD140a (BioLegend, #APA5), CD90.2 (BioLegend, #30-H12), TER119 (BioLegend, #116234), CD326 (Thermo Fisher Scientific, #50-163-76), CD45 (BioLegend, #103126), CD31 (BioLegend, #102422), CD202b (Thermo Fisher Scientific, #15-5987-82), Brilliant Violet 421 Streptavidin (BioLegend, #405226) and DAPI staining solution (Thermo Fisher Scientific, # 62248).

For Single cell RNA-seq of young and old wounds using 10x Genomics Chromium single-cell
Single cell RNA-sequencing was performed on all live PDGFRα+ Lin- (TER119-/CD326-/CD202b-/CD45-/CD31-) cells in the wounded area from young and old mice, 7 days after wounding. To this end, we pooled cells from 10 young (3-4 months), or 10 old (25-27 months) male C57BL/6 mouse from the NIA aged colony. FACS sorting was performed as described above, with the exception that all PDGFRα+ Lin- (TER119-/CD326-/CD202b-/CD45-/CD31-) cells were sorted into fibroblast growth media. Cells were then spun down at 300xg for 5 minutes at 4oC and resuspended in fibroblast growth media at a concentration of 300 cells/µL. In total, 18,000 young cells and 18,000 old cells were loaded onto a respective 10x Genomics Chromium chip per

| | |
|---|---|
| | manufacturer's recommendations.<br><br>For Single cell RNA-seq of fast-healing and slow-healing old wounds using 10x Genomics Chromium single-cell<br>Single cell RNA-sequencing was performed on all live cells in the wounded area from 2 fast-healing and 2 slow-healing old mice, 7 days after wounding. Live/dead staining was performed using 1µg/mL propidium iodide (BioLegend). FACS sorting was performed on a BD FACS Aria Fusion sorter using a 100µm nozzle. Cells were sorted into chilled fibroblast growth media. Cells were then spun down at 300xg for 5 min at 4oC and resuspended in fibroblast growth media at a concentration of 1000-1500 cells/µL. Cells were loaded onto a 10x Genomics Chromium chip as described above. |
| Instrument | FACS analysis was performed on an LSR II flow cytometer (BD Biosciences), and FACS sorting was performed on a BD FACS Aria II sorter or a BD FACS Fusion sorter, using a 100 µm nozzle. All instruments were housed in the Stanford Shared FACS Facility. |
| Software | All flow cytometry data was analyzed using FlowJo version 10.0.7. |
| Cell population abundance | To determine the purity of the primary fibroblasts from young and old mice, FACS analysis was performed on fibroblast cultures at passage 3. Cultures were stained for PDGFRα (a general fibroblast marker), in combination with marker genes for possible contaminants, including B and T cells, granulocytes, monocytes, and dendritic cells. Our FACS analysis of fibroblasts from fresh tissues gates out possible contaminants, including immune cells (CD45), endothelial cells (CD31, TIE2), epithelial (EpCAM) and red blood cells (TER119), and gates for PDGFRα+ cells that are either THY1+ or THY1-. In the context of single-cell RNAseq from young and old wounds, we gated for CD45-/CD31-/TIE2-/EpCAM-/TER119/PDGFRα+ cells. |
| Gating strategy | Gating was determined using fluorescent-minus-one controls for each color used in each FACS experiment to ensure that positive populations were solely associated with the antibody for that specific marker (see Extended Data Fig. 10). |

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

# A glucose-sensing neuron pair regulates insulin and glucagon in *Drosophila*

https://doi.org/10.1038/s41586-019-1675-4

Received: 26 September 2018

Accepted: 16 September 2019

Published online: 23 October 2019

Yangkyun Oh[1], Jason Sih-Yu Lai[1,2], Holly J. Mills[1,3], Hediye Erdjument-Bromage[1], Benno Giammarinaro[1,4], Khalil Saadipour[1], Justin G. Wang[1,5], Farhan Abu[1,6], Thomas A. Neubert[1] & Greg S. B. Suh[1,7]*

Although glucose-sensing neurons were identified more than 50 years ago, the physiological role of glucose sensing in metazoans remains unclear. Here we identify a pair of glucose-sensing neurons with bifurcated axons in the brain of *Drosophila*. One axon branch projects to insulin-producing cells to trigger the release of *Drosophila* insulin-like peptide 2 (dilp2) and the other extends to adipokinetic hormone (AKH)–producing cells to inhibit secretion of AKH, the fly analogue of glucagon. These axonal branches undergo synaptic remodelling in response to changes in their internal energy status. Silencing of these glucose-sensing neurons largely disabled the response of insulin-producing cells to glucose and dilp2 secretion, disinhibited AKH secretion in corpora cardiaca and caused hyperglycaemia, a hallmark feature of diabetes mellitus. We propose that these glucose-sensing neurons maintain glucose homeostasis by promoting the secretion of dilp2 and suppressing the release of AKH when haemolymph glucose levels are high.

Glucose-sensing neurons respond to glucose or its metabolites, which act as signalling cues to regulate their neuronal activity. According to the glucostatic hypothesis proposed in 1953, feeding and related behaviours are regulated by neurons in the brain that sense changes in glucose levels in the blood[1]. Despite the discovery of glucose-sensing neurons in the hypothalamus through electrophysiological methods more than ten years later[2], the physiological role of these neurons remained unclear[3,4] until recently, when a population of glucose-excited neurons in the *Drosophila* brain were determined to function as an internal nutrient sensor to mediate the animal's consumption of sugar[5]. A large number of glucose-sensing neurons appear to be present in animals[6]; we speculated that these neurons mediate physiological functions that are critical for the wellbeing of the animal, including glucose homeostasis. Here we report the identification of a pair of glucose-excited neurons in the *Drosophila* brain that maintain glucose homeostasis by coordinating the activity of the two key hormones involved in the process: insulin and glucagon.

## CN neurons project to the PI and CC

To identify neurons that respond to sugar on the basis of its nutritional value, we used a two-choice assay[7] to screen Vienna tiles (VT)-Gal4 *Drosophila* lines[8] that had been crossed to *UAS-Kir2.1, tub-Gal80$^{ts}$* flies (inward-rectifier potassium ion channel allele *Kir2.1* with tubulin-temperature-sensitive Gal80) for defects in their ability to select nutritive D-glucose over non-nutritive L-glucose (Extended Data Fig. 1a, see Methods). We isolated two independent Gal4 lines, *VT58471* and *VT43147-Gal4*, that failed to select D-glucose after periods of starvation and appeared to contain dorsolateral cells that resemble those that are labelled by the corazonin (*Crz*)-*Gal4* line[9] (Extended Data Fig. 1b, c, arrowheads). Flies in which *Crz-Gal4*-expressing neurons had been inactivated failed to select D-glucose even when starved (Extended Data Fig. 1d). These results suggest that the dorsolateral neurons labelled by *Crz-Gal4* and two candidate Gal4 lines mediate the behavioural response to sugar.

We used a Crz antibody to confirm the identity of the dorsolateral neurons (Fig. 1a, top right). A previous study demonstrated that a subset of Crz-expressing neurons also express short neuropeptide F (sNPF)[10]. Immunolabelling revealed that the dorsolateral neurons expressing Crz indeed express sNPF (Fig. 1a, bottom right). On the basis of these findings, we named these Crz$^+$sNPF$^+$ neurons CN neurons. To restrict Gal4 expression to a few cells that include the dorsolateral neurons, we crossed *VT58471-Gal4* to choline acetyltransferase (*ChAT*)-*Gal80*, generating *CN-Gal4*, which unambiguously labelled a pair of CN neurons when crossed to *UAS-mCD8::GFP* (Fig. 1a, left). Flies in which these dorsolateral neurons were inactivated using *CN-Gal4* failed to select D-glucose when starved (Fig. 1b). Each CN cell body projects an axon that bifurcates to form two major branches (Fig. 1a). One branch (axon 1) projects to the pars intercerebralis (PI) region of the brain and the other branch (axon 2) projects ventrolaterally towards the corpora cardiaca (CC)[11,12] (Fig. 1a, c, Extended Data Fig. 2a). We used an intersectional approach to define these projections further, thereby validating that axon 1 innervates the PI and axon 2 projects to the CC (Fig. 1d, Extended Data Fig. 2b, c, see Methods). We also used this approach to induce the expression of tetanus toxin (TNT)[13] to silence a pair of CN neurons. These flies failed to choose D-glucose even after starvation when CN neurons were inactivated (Fig. 1e, see Methods). This provided further evidence of the contribution of the pair of the dorsolateral CN neurons to glucose-evoked behaviour.

[1]Skirball Institute of Biomolecular Medicine, Department of Cell Biology, Neuroscience Institute, New York University School of Medicine, New York, NY, USA. [2]QPS-Qualitix Taiwan, Ren-Ai Road, Taipei, Taiwan. [3]Ascend Public Charter Schools, New York, NY, USA. [4]Vision Sciences Graduate Program, School of Optometry, UC Berkeley, Berkeley, CA, USA. [5]Salk Institute for Biological Studies, La Jolla, CA, USA. [6]Department of Genetics and Development, Columbia University, New York, NY, USA. [7]Department of Biological Sciences, Korea Advanced Institute of Science and Technology, Daejeon, South Korea. *e-mail: seongbaesuh@kaist.ac.kr

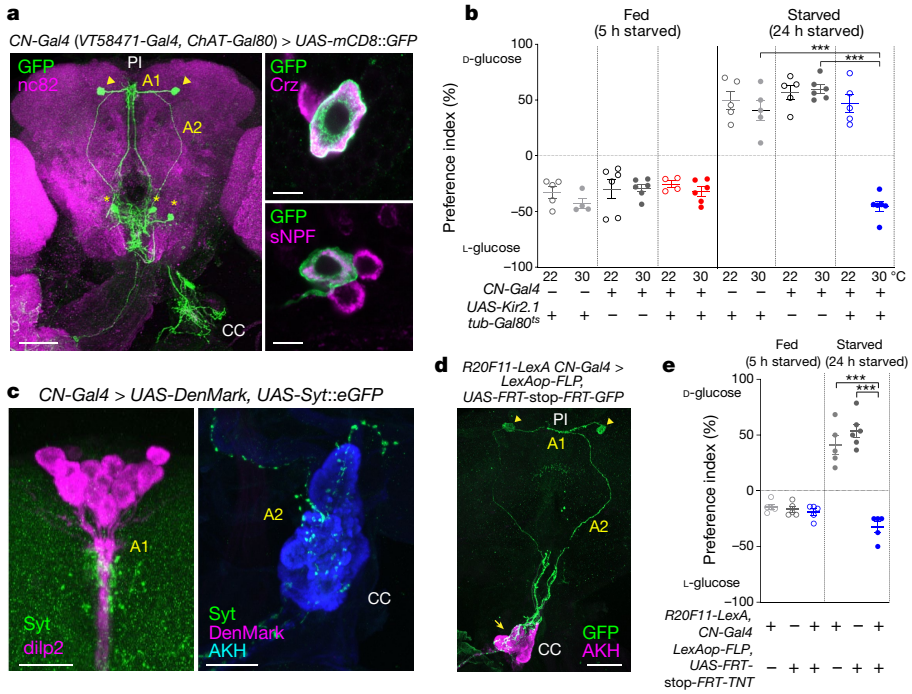



**Fig. 1 | A pair of glucose-sensing CN neurons in the brain show a unique projection pattern. a**, CN neurons labelled with GFP (green) counterstained with nc82 antibody (magenta) in the brain and CC. CN neurons (arrowheads) extend their neurites centrally (A1) and ventrolaterally (A2). Asterisks denote unrelated cells labelled by *CN-Gal4*. Scale bar, 50 µm. CN cell bodies labelled with GFP (green) co-stained with Crz (magenta, top right) and sNPF antibodies (magenta, bottom right). Scale bar, 5 µm. **b**, Inactivation of CN neurons by expressing *UAS-Kir2.1* and *tub-Gal80$^{ts}$* under the control of *CN-Gal4* at 30 °C blunts a preference for D-glucose in starved flies. **c**, CN axons innervate insulin-producing cells (IPCs), stained with dilp2 antibody (magenta), via axon 1 (A1, left) and AKH-producing cells, stained with AKH antibody (cyan), via axon 2 (A2, right). CN axons and dendrites are stained with GFP (green) and DenMark detectable dsRed (magenta, see Methods) antibodies, respectively. Scale bar, 20 µm. **d**, Intersectional labelling of CN neurons by GFP (green) co-stained with AKH antibody (magenta). Scale bar, 50 µm. **e**, Intersectional silencing of CN neurons blunts a preference for D-glucose in starved flies. Images shown are *z*-stacked projections, except in **a**, right. In all figures, plots show mean ± s.e.m. ***$P$ < 0.001; one-way ANOVA with Tukey post hoc test. Sample sizes and statistical analyses are shown in Supplementary Table 1.

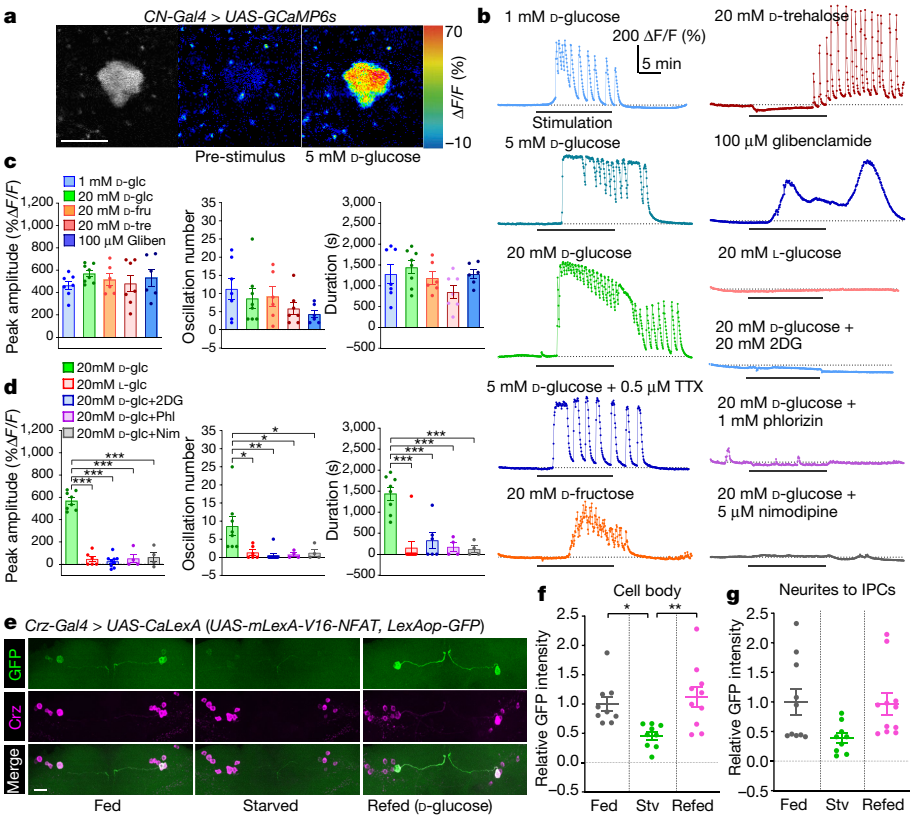



**Fig. 2 | CN neurons are activated by nutritive sugars, but not by non-nutritive sugars. a**, CN neurons expressing GCaMP before (middle) and after (right) D-glucose application. Scale bar, 5 µm. **b**–**d**, Representative traces (**b**) and quantifications (**c**, **d**) of calcium responses of CN neurons to D-glucose (D-glc), D-fructose (D-fru), D-trehalose (D-tre) and glibenclamide (Gliben) (**c**) and L-glucose (L-glc) and D-glucose mixed with 2-D-deoxyglucose (2DG), phlorizin (Phl) or nimodipine (Nim) (**d**). **e**–**g**, Representative images revealed by native CaLexA-driven GFP with anti-Crz staining (**e**) and quantifications (**f**, **g**) of GFP intensity from CN cell bodies (**f**) and neurites to IPCs (**g**) of fed, starved (Stv) or refed flies carrying *Crz-Gal4* and *UAS-CaLexA*; see Methods. Scale bar, 20 µm. Images are *z*-stacked projections. *$P$ < 0.05, **$P$ < 0.01 and ***$P$ < 0.001; one-way ANOVA with Tukey post hoc test. Sample sizes and statistical analyses are shown in Supplementary Table 1.

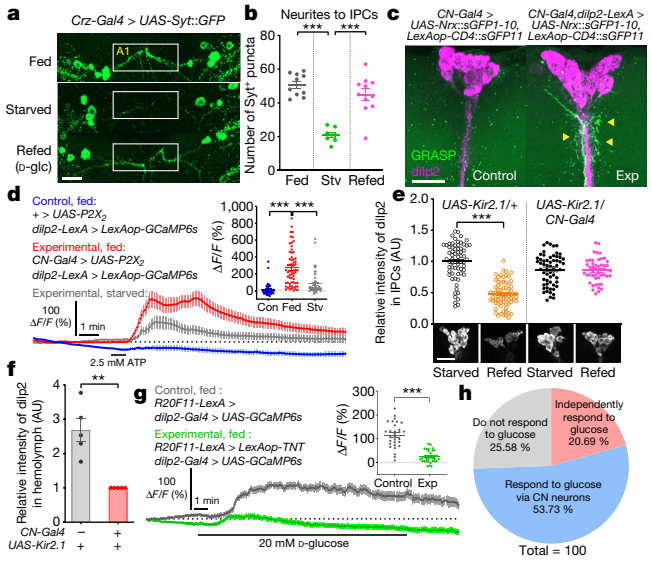



**Fig. 3 | IPC activity and dilp2 secretion require an excitatory signal from CN neurons. a**, **b**, Representative images (**a**) and quantification of the number of Syt–GFP$^+$ puncta (green) (**b**) in axon 1 (A1) of CN neurons in fed, starved or refed flies carrying *Crz-Gal4* and *UAS-Syt::GFP*. The outlined region shows CN axonal projections to IPCs. **c**, Native GRASP-induced fluorescence (green, arrowheads), co-stained with dilp2 antibody (magenta). **d**, Average GCaMP traces and Δ*F*/*F* (max) quantifications from IPCs of fed or starved flies in which CN neurons were stimulated, or those of fed control (Con) flies. **e**, **f**, Relative intensities of dilp2 immunoreactivity in IPCs (**e**) and tryptic peptide of dilp2 B chain from haemolymph (**f**) of fed flies in which CN neurons were inactivated, or those of control flies. **g**, Average GCaMP traces and Δ*F*/*F* (max) quantifications from IPCs of fed flies in which CN neurons were silenced by TNT in response to D-glucose or those of control flies. **h**, IPCs comprise three subpopulations according to their response to glucose; see Methods. Scale bars, 20 µm. Images are *z*-stacked projections. $^{**}P < 0.01$ and $^{***}P < 0.001$; one-way ANOVA with Tukey post hoc test for (**b**, **d**) and unpaired two-tailed *t*-test for (**e**, **f**, **g**). See Supplementary Table 1 for the sample sizes and statistical analyses. AU, arbitrary units.

## CN neurons are glucose-excited

We next sought to determine whether CN neurons respond to glucose and other sugars. Calcium-imaging studies using ex vivo brain preparations of flies carrying the calcium indicator *UAS-GCaMP6s*[14] and *CN-Gal4* revealed that CN neurons were robustly activated by D-glucose with substantial calcium oscillations (Fig. 2a–c, Extended Data Fig. 3a–e). CN neurons also responded to D-trehalose and D-fructose, which are found in the haemolymph, but failed to respond to (1) the non-nutritive sugar L-glucose (Fig. 2b, d); (2) the non-haemolymph sugar sucrose; and (3) the non-sugar nutrients amino acids (Extended Data Fig. 3f–j). D-Glucose and D-trehalose are key sugars in the haemolymph, although D-trehalose stimulates the activity of CN neurons only after a substantial delay (about 12 min), possibly because it requires additional metabolic steps to be converted to glucose. D-Fructose applied at 20 mM activated CN neurons, although the concentration of D-fructose in the haemolymph is much lower (<2 mM)[15]. These findings suggest that the pair of CN neurons responds only to D-glucose under normal physiological conditions.

We next determined whether activation of CN neurons by D-glucose requires glucose metabolism inside the cell. Exposing the brain to D-glucose mixed with 2DG, phlorizin or nimodipine, which inhibits glycolysis, glucose transport or voltage-gated calcium channels, respectively, blunted the glucose-induced stimulation of CN neurons (Fig. 2b, d). In the presence of pyruvate (an end product of glycolysis), the CN neurons demonstrated activity similar to that seen in the presence of other haemolymph sugars (Extended Data Fig. 3f–j). Application of the ATP-sensitive potassium channel ($K_{ATP}$)[16] blocker glibenclamide[17] resulted in activation of CN neurons (Fig. 2b, c). Furthermore, glucose-induced calcium transients of these neurons were not abrogated by the application of the sodium-channel blocker tetrodotoxin (TTX) (Fig. 2b). Using RNA-mediated interference (RNAi) lines, we also determined that glucose transporter 1 (*Glut1*), hexokinase C (*Hex-C*), a subunit of the $K_{ATP}$ channel (*SUR1*) and the voltage-gated calcium channel are required in CN neurons for the two-choice behaviour (Extended Data Fig. 4a–c). Consistent with the behavioural results, the glucose-induced calcium response of CN neurons requires *Glut1*, *SUR1* and a voltage-gated calcium channel (Extended Data Fig. 4d–h), further supporting the role of the intracellular glucose metabolic pathway in stimulating CN neuronal activity.

We next used the calcium-dependent nuclear import of LexA (CaLexA) system[18] to measure cellular activity in CN neurons in intact flies, and found that GFP signal driven by the CaLexA system in starved flies was significantly reduced compared to the signal in fed flies, and the signal was restored when starved flies were refed D-glucose (Fig. 2e–g, Extended Data Fig. 4i, j). These results suggest that the activity of CN neurons is stimulated by the increase in glucose levels observed under fed conditions. In addition to the altered CaLexA signals, we evaluated the effect of glucose on the number and intensity of synaptotagmin (Syt)[19]–GPF$^+$ puncta in fed, starved and refed animals. The Syt–GPF$^+$ signals decreased significantly in axon 1 in starved animals and returned to normal levels after the flies were fed with D-glucose (Fig. 3a, b, Extended Data Fig. 5a, c). However, this nutrient-dependent plasticity was not observed in *Crz-Gal4*-labelled axonal processes that did not originate from the dorsolateral CN neurons (Extended Data Fig. 5e–h).

## CN neurons promote dilp2 release

We next sought to determine whether CN neurons are coupled with IPCs[20] at the synaptic level. We used a modified GFP reconstitution across synaptic partners (GRASP) method[21], and found that the GRASP signals were visible around the synapse between CN neurons and insulin-producing cells (IPCs) (Fig. 3c), indicating physical coupling between CN neurons and IPCs.

To determine whether the coupling between CN neurons and IPCs is functional, we expressed ATP-gated $P2X_2$ purine receptors[22] in CN neurons and the calcium indicator GCaMP6s[14] in IPCs, and then stimulated CN neurons using ATP while recording from the IPCs. As shown in Fig. 3d, ATP-induced CN-neuron activity was accompanied by a significant increase in the amplitude of GCaMP signals in the IPCs in fed flies; this effect was reduced in starved flies (Fig. 3d, Extended Data Fig. 6a, b). This finding supports the hypothesis that the nutrient-dependent synaptic changes observed between CN neurons and IPCs have functional consequences. The CN neurons did not appear to be functionally coupled to glucose-excited diuretic hormone 44 (Dh44) neurons[5] (Extended Data Fig. 6c–f). Furthermore, we investigated whether CN neuronal activity is required for dilp2 secretion from IPCs. We observed a significant reduction in the intensity of dilp2 immunoreactivity in the IPCs of fed control flies, but not in fed flies in which CN neurons had been inactivated (Fig. 3e). These results suggest that an excitatory signal from the CN neurons contributes to the secretion of dilp2 from IPCs in response to increased glucose levels. Using mass spectrometry and dot blot assay, we further validated that the flies carrying *CN-Gal4* and *UAS-Kir2.1* had lower dilp2 levels circulating in the haemolymph than wild-type flies, in contrast to the higher dilp2 levels found in IPCs (Fig. 3f, Extended Data Fig. 7a–c, e, f).

To further clarify the role of CN neurons in mediating glucose-evoked activity in IPCs, we inactivated CN neurons by expressing TNT[13] and then examined the responsiveness of IPCs to glucose. The amplitude of calcium signals in IPCs that had been exposed to D-glucose was significantly reduced when the CN neurons were inactivated (Fig. 3g). Furthermore, we found that IPCs harbour at least three subpopulations of neurons with distinct responses to glucose or $K_{ATP}$ channel blocker (Fig. 3h and Extended Data Fig. 8a–f). These findings suggest that CN

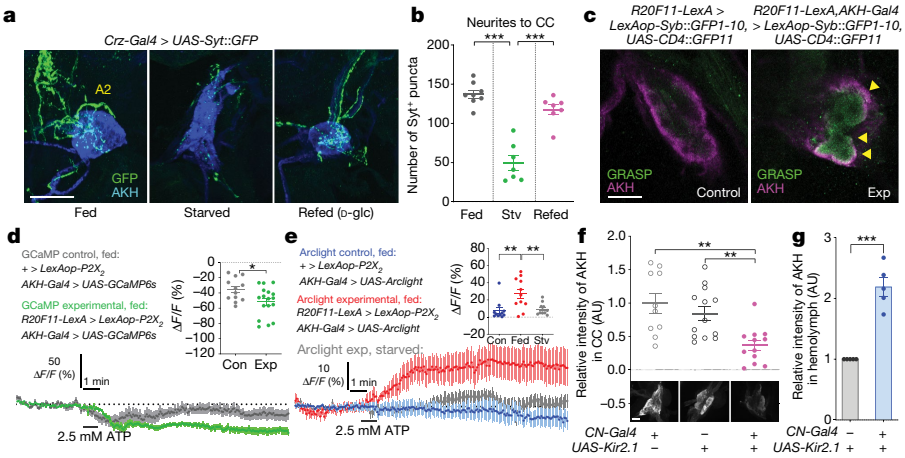



**Fig. 4 | AKH retention in AKH-producing cells requires an inhibitory signal from CN neurons. a**, **b**, Representatives images stained with GFP and AKH antibodies (**a**) and quantifications of the number of Syt–GFP$^+$ puncta (green) (**b**) in axon 2 (A2) of CN neurons in fed, starved or refed flies carrying *Crz-Gal4* and *UAS-Syt::GFP*. Scale bar, 50 µm. **c**, Synaptic GRASP-induced GFP signals (arrowheads) co-stained with AKH antibody (magenta). Scale bar, 20 µm. **d**, **e**, Average GCaMP traces and Δ*F*/*F* (min) quantifications (**d**) and average Arclight traces and Δ*F*/*F* (max) quantifications (**e**) from AKH-producing cells of fed or starved flies in which CN neurons were activated or those of control flies. **f**, **g**, Relative intensities of AKH immunoreactivity in AKH-producing cells (**f**) and tryptic peptide of AKH from haemolymph (**g**) of fed flies in which CN neurons were inactivated or those of control flies. Scale bar, 20 µm. Images are *z*-stacked projections except in **c**. $^*P < 0.05$, $^{**}P < 0.01$ and $^{***}P < 0.001$; one-way ANOVA with Tukey post hoc test (**b**, **e**, **f**) and unpaired two-tailed *t*-test (**d**, **g**). Sample sizes and statistical analyses are shown in Supplementary Table 1.

neuronal activity is required for the majority of IPCs to respond to glucose.

## CN neurons inhibit AKH secretion from CC

To determine whether nutrient-dependent plasticity also occurs in axon 2 of the CN neurons, we monitored the number and intensity of Syt–GFP$^+$ puncta before and after feeding flies with D-glucose. We observed a significant reduction in these parameters in starved flies and a restoration to normal levels after refeeding starved flies with D-glucose (Fig. 4a, b, Extended Data Fig. 5b, d). This raised the possibility of coupling between CN neurons and AKH-producing cells. Using a modified GRASP method[23], we observed GRASP fluorescent signals around AKH-producing cells (Fig. 4c). To determine whether there is any functional connectivity between these cells, we activated the CN neurons while monitoring the activity of AKH-producing cells and found that calcium transients in the AKH-producing cells appeared to decrease during activation of CN neurons (Fig. 4d).

To probe this observation further, we expressed the Arclight receptor[24], which increases fluorescent signals when cells become hyperpolarized, in AKH-producing cells and $P2X_2$ receptors in CN neurons. When the CN neurons were activated using ATP, the Arclight fluorescence intensity in fed flies increased significantly compared with that in starved flies (Fig. 4e), validating the occurrence of nutrient-dependent changes in the synapses between CN neurons and AKH-producing cells. Notably, when CN neurons were inactivated, the intracellular AKH levels decreased significantly compared with controls (Fig. 4f). Using mass spectrometry and dot blot assay, we confirmed significantly higher levels of AKH in haemolymph of flies carrying *CN-Gal4* and *UAS-Kir2.1* compared with those in control flies (Fig. 4g, Extended Data Fig. 7d, g, h). These findings suggest that CN neuronal activity inhibits the release of AKH from the CC and the increase of AKH levels in haemolymph.

## sNPF is the functional neurotransmitter

We next investigated the identities of the key neurotransmitters in axon 1 and axon 2 for regulating the functionally opposing synaptic activities. We tested the role of Crz and sNPF in the two-choice behaviour using RNAi lines and found that sNPF in CN neurons and sNPF receptor in the postsynaptic IPCs, but not Crz or its receptor, are important (Fig. 5a, b, Extended Data Fig. 9a–d). We also found that sNPF but not Crz levels in CN neurons were significantly reduced when CN neurons were exposed to D-glucose (Fig. 5c, Extended Data Fig. 9e). We observed that approximately a half of the IPCs that had responded to glucose failed to respond glucose when the dominant-negative allele of *sNPF* receptor[25] was expressed in IPCs (Fig. 5d and Extended Data Fig. 8g–i). Furthermore, we observed that intracellular AKH levels remained high in AKH-producing cells in fed control flies, but declined significantly in fed flies in which the function of *sNPF* receptor[25] was inhibited in AKH-producing cells (Fig. 5e).

Finally, we determined whether sNPF alters activity of IPCs and/or the CC. The activity of IPCs was significantly stimulated by the application of sNPF[26] (Extended Data Fig. 10a, b), whereas CC activity was significantly inhibited by sNPF (Fig. 5f). These functionally opposing effects of sNPF are probably mediated by $G_q$ in IPCs and by $G_{i/o}$ in AKH-producing cells via the sNPF receptor, which is a G-protein-coupled receptor[27]. Exposing the brain to U73122, a PLC inhibitor that inhibits the $G_q$ pathway, eliminated the glucose-evoked activation of IPCs, but had no effect on sNPF-induced inhibition of AKH-expressing cells (Fig. 5f, g). Conversely, exposing the brain to pertussis toxin, a $G_i$ inhibitor, blunted the sNPF-induced inhibition of AKH-producing cells, but had no effect on the glucose-evoked activation of IPCs (Fig. 5f, g). These results indicate that axon 1 and axon 2 can have opposing synaptic activities through a mechanism involving the same neurotransmitter and receptor but with distinct downstream factors coupled with opposing outputs.

To determine whether CN neuronal activity can alter circulating sugar levels in flies, we monitored circulating concentrations of glucose and trehalose in haemolymph, and found that they were significantly increased in flies in which CN neurons were inactivated compared with controls (Fig. 5h, Extended Data Fig. 10c). This finding illustrates that dysfunctional CN neuronal input to IPCs and AKH-producing cells results in a defect in glucose homeostasis.

## Discussion

We identified and characterized a pair of glucose-sensing neurons in the *Drosophila* brain that have an essential role in maintaining glucose homeostasis. This was achieved by counterbalancing the activities of *Drosophila* equivalents of insulin- and glucagon-producing cells. When

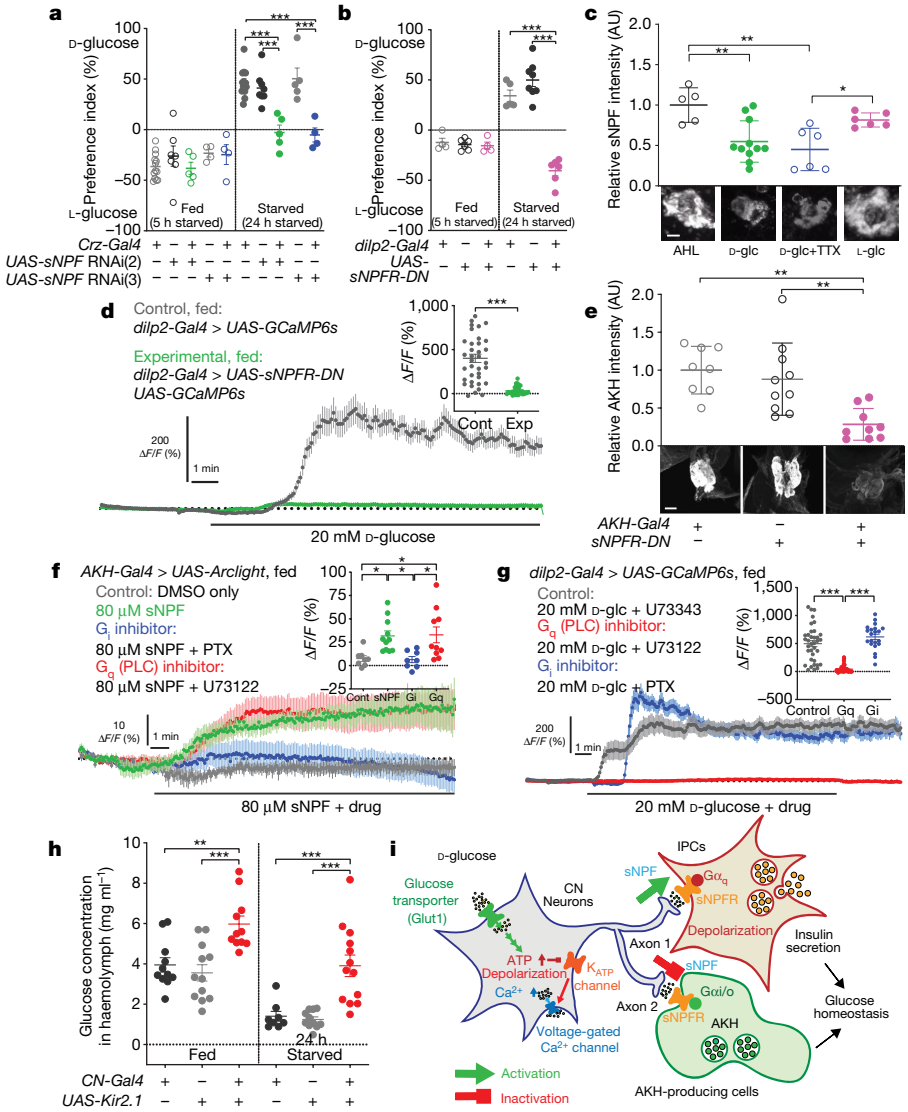



**Fig. 5 | sNPF is the functional neurotransmitter of CN neurons. a**, Knockdown of *sNPF* in flies carrying *Crz-Gal4* and *UAS-sNPF* RNAi abolishes a preference for D-glucose in starved flies. **b**, Expression of the dominant-negative sNPF receptor (*UAS-sNPFR-DN*) using *dilp2-Gal4* blunts a preference for D-glucose in starved flies. **c**, Immunoreactivity of intracellular sNPF in CN neurons when the brains were incubated in 80 mM sucrose, D-glucose, D-glucose mixed with 0.5 μM TTX (D-Glc + TTX) or L-glucose in artificial haemolymph-like solution (AHL). Scale bar, 5 μm. **d**, Average GCaMP traces and Δ*F*/*F* (max) quantifications from IPCs of fed flies in which *UAS-sNPFR-DN* was expressed in IPCs in response to D-glucose, or those of control flies. **e**, Immunoreactivity of intracellular AKH in CC of flies carrying *UAS-sNPFR-DN* and *AKH-Gal4*, or those of control flies. Scale bar, 20 μm. **f**, Average Arclight traces and Δ*F*/*F* (max) quantifications from CC in response to sNPF mixed with $G_i$ inhibitor, pertussis toxin (PTX), $G_q$ (phospholipase C, PLC) inhibitor U73122 or DMSO-only controls. **g**, Average GCaMP traces and Δ*F*/*F* (max) quantifications from IPCs in response to D-glucose mixed with U73122, PTX or U73343, a non-functional enantiomer of U73122; see Methods. **h**, Circulating glucose levels of flies in which CN neurons were silenced, or those in control flies. **i**, Schematics of the functional connectivity among CN neurons, IPCs and AKH-producing cells. Images are *z*-stacked projections. **P* < 0.05, ***P* < 0.01 and ****P* < 0.001; one-way ANOVA with Tukey post hoc test and unpaired two-tailed *t*-test in **d**. Sample sizes and statistical analyses are shown in Supplementary Table 1.

food consumption leads to a rise in haemolymph sugar levels, CN neurons excite the IPCs through sNPF and its receptor, which appear to be coupled to the $G_q$ signalling cascade to induce the secretion of dilp2, while suppressing the release of AKH by using the same sNPF receptor, which in this case is coupled with $G_i$ signalling pathway (Fig. 5i, Extended Data Fig. 10d). We speculate that precise control of these opposing functions is facilitated because the nutrient-dependent plastic changes arise from a single cell.

This study demonstrates how the activity of the two key endocrine systems is coordinated in metazoans and that their coordination is under the direct control of glucose-sensing neurons. Such coordination has been proposed to occur in mammals via the sympathetic and parasympathetic nerves that connect the pancreatic islets with glucose-sensing neurons in the hypothalamus and hindbrain[28]. The finding that a large proportion of IPCs respond to glucose through CN neurons in insects raises an intriguing possibility that both direct and indirect mechanisms control endocrine function in mammals[29,30]. Finally, this work may shed light on the function of glucose-sensing neurons. Further research is needed to understand how these regulatory processes are affected by excessive nutrition and other metabolic disturbances, including obesity.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information,

acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1675-4.


1. Mayer, J. Glucostatic mechanism of regulation of food intake. *N. Engl. J. Med.* **249**, 13–16 (1953).
2. Oomura, Y. et al. Reciprocal activities of the ventromedial and lateral hypothalamic areas of cats. *Science* **143**, 484–485 (1964).
3. Parton, L. E. et al. Glucose sensing by POMC neurons regulates glucose homeostasis and is impaired in obesity. *Nature* **449**, 228–232 (2007).
4. Levin, B. E. Neuronal glucose sensing: still a physiological orphan? *Cell Metab.* **6**, 252–254 (2007).
5. Dus, M. et al. Nutrient sensor in the brain directs the action of the brain–gut axis in *Drosophila*. *Neuron* **87**, 139–151 (2015).
6. López-Gambero, A. J., Martínez, F., Salazar, K., Cifuentes, M. & Nualart, F. Brain glucose-sensing mechanism and energy homeostasis. *Mol. Neurobiol.* **56**, 769–796 (2019).
7. Dus, M., Min, S., Keene, A. C., Lee, G. Y. & Suh, G. S. B. Taste-independent detection of the caloric content of sugar in *Drosophila*. *Proc. Natl Acad. Sci. USA* **108**, 11644–11649 (2011).
8. Pfeiffer, B. D. et al. Tools for neuroanatomy and neurogenetics in *Drosophila*. *Proc. Natl Acad. Sci. USA* **105**, 9715–9720 (2008).
9. Tayler, T. D., Pacheco, D. A., Hergarden, A. C., Murthy, M. & Anderson, D. J. A neuropeptide circuit that coordinates sperm transfer and copulation duration in *Drosophila*. *Proc. Natl Acad. Sci. USA* **109**, 20697–20702 (2012).
10. Kapan, N., Lushchak, O. V., Luo, J. & Nässel, D. R. Identified peptidergic neurons in the *Drosophila* brain regulate insulin-producing cells, stress responses and metabolism by coexpressed short neuropeptide F and corazonin. *Cell. Mol. Life Sci.* **69**, 4051–4066 (2012).
11. Kim, S. K. & Rulifson, E. J. Conserved mechanisms of glucose sensing and regulation by *Drosophila* corpora cardiaca cells. *Nature* **431**, 316–320 (2004).
12. Lee, G. & Park, J. H. Hemolymph sugar homeostasis and starvation-induced hyperactivity affected by genetic manipulations of the adipokinetic hormone-encoding gene in *Drosophila* melanogaster. *Genetics* **167**, 311–323 (2004).
13. Sweeney, S. T., Broadie, K., Keane, J., Niemann, H. & O'Kane, C. J. Targeted expression of tetanus toxin light chain in *Drosophila* specifically eliminates synaptic transmission and causes behavioral defects. *Neuron* **14**, 341–351 (1995).
14. Chen, T.-W. et al. Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).
15. Miyamoto, T., Slone, J., Song, X. & Amrein, H. A fructose receptor functions as a nutrient sensor in the *Drosophila* brain. *Cell* **151**, 1113–1125 (2012).
16. Ashcroft, F. M. & Rorsman, P. $K_{ATP}$ channels and islet hormone secretion: new insights and controversies. *Nat. Rev. Endocrinol.* **9**, 660–669 (2013).
17. Bryan, J. & Aguilar-Bryan, L. The ABCs of ATP-sensitive potassium channels: more pieces of the puzzle. *Curr. Opin. Cell Biol.* **9**, 553–559 (1997).
18. Masuyama, K., Zhang, Y., Rao, Y. & Wang, J. W. Mapping neural circuits with activity-dependent nuclear import of a transcription factor. *J. Neurogenet.* **26**, 89–102 (2012).
19. Zhang, Y. Q., Rodesch, C. K. & Broadie, K. Living synaptic vesicle marker: synaptotagmin–GFP. *Genesis* **34**, 142–145 (2002).
20. Rulifson, E. J., Kim, S. K. & Nusse, R. Ablation of insulin-producing neurons in flies: growth and diabetic phenotypes. *Science* **296**, 1118–1120 (2002).
21. Feinberg, E. H. et al. GFP reconstitution across synaptic partners (GRASP) defines cell contacts and synapses in living nervous systems. *Neuron* **57**, 353–363 (2008).
22. Lima, S. Q. & Miesenböck, G. Remote control of behavior through genetically targeted photostimulation of neurons. *Cell* **121**, 141–152 (2005).
23. Macpherson, L. J. et al. Dynamic labelling of neural connections in multiple colours by trans-synaptic fluorescence complementation. *Nat. Comm.* **6**, 10024 (2015).
24. Cao, G. et al. Genetically targeted optical electrophysiology in intact neural circuits. *Cell* **154**, 904–913 (2013).
25. Lee, K.-S. et al. *Drosophila* short neuropeptide F signalling regulates growth by ERK-mediated insulin signalling. *Nat. Cell Biol.* **10**, 468–475 (2008).
26. Shang, Y. et al. Short neuropeptide F is a sleep-promoting inhibitory modulator. *Neuron* **80**, 171–183 (2013).
27. Garczynski, S. F., Brown, M. R., Shen, P., Murray, T. F. & Crim, J. W. Characterization of a functional neuropeptide F receptor from *Drosophila* melanogaster. *Peptides* **23**, 773–780 (2002).
28. Thorens, B. Neural regulation of pancreatic islet cell mass and function. *Diabetes Obes. Metab.* **16** (Suppl 1), 87–95 (2014).
29. Pipeleers, D. G. Heterogeneity in pancreatic β-cell population. *Diabetes* **41**, 777–781 (1992).
30. Bonner-Weir, S. & Aguayo-Mazzucato, C. Pancreatic β-cell heterogeneity revisited. *Nature* **535**, 365–366 (2016).

# Methods

### Fly strains

Flies were raised on standard cornmeal-molasses food (940 ml water, 9 g agar, 15 g yeast, 36 g cornmeal, 36 ml molasses, 1.12 g tegosept, 3.8 ml propionic acid, total 1 l of fly food) at 23 °C with 12 h:12 h light:dark cycles. Fly strains were obtained as described in Supplementary Table 2. All the lines used for behavioural testing were backcrossed into $w^{1118}$ (Bloomington no. 6326) background for at least five generations.

### Two-choice assay

The two-choice preference assay was performed as previously described[7]. In brief, approximately 30–40 male flies (1–3 days old) were collected under $CO_2$ anaesthesia and allowed to recover for 2–3 more days. The flies were starved in an empty vial with wetted Kimwipe with 2–3 ml of distilled water for 5 h (fed condition) or 24 h (starved condition). They were introduced into a two-choice arena containing two food sources, 50 mM D-glucose and 200 mM L-glucose, that were colored with a tasteless food dye (1% of red food dye and 0.7% of green food dye, McCORMICK) where flies were allowed to feed for 2 h in the dark. The majority of flies preferred sweeter L-glucose (200 mM) when fed, but chose nutritive D-glucose (50 mM) when starved[5,7]. For experiments using *UAS-Kir2.1, tub-Gal80$^{ts}$*, flies were raised at 18 °C before transferring to 30 °C for 48–72 h to inactivate Gal80 in order to express *Kir2.1*. Following behavioural testing was conducted at 23 °C. For experiments using flies that do not bear *tubulin-Gal80$^{ts}$*, behavioural testing was conducted at 23 °C. All sugars (D-glucose, D-trehalose and D-fructose) were purchased from Sigma except L-glucose (Carbosynth). Statistical analyses were performed using GraphPad Prism 8.1.1. Food preference was scored as a percentage of preference index (% PI), calculated as shown below (where $f_n$ is the number of flies that ate food $n$):

$$\%\ \mathrm{PI} = \frac{(f_1 + 0.5 \times (f_1 + f_2)) - (f_2 + 0.5 \times (f_1 + f_2))}{f_1 + f_2} \times 100\%$$

### Two-choice behavioural screen using VT-Gal4 drives

We obtained more than 100 Gal4 drivers from the Vienna *Drosophila* RNAi Center (VDRC, Vienna Tiles (VT) library)[8]. We used the two-choice assay to screen 95 VT-Gal4 drivers, which had been crossed to *UAS-Kir2.1, tub-Gal80$^{ts}$* (inward-rectifier potassium ion channel allele, *Kir2.1*, with tubulin-temperature-sensitive Gal80). We found 18 candidate Gal4 drivers that preferred L-glucose even when they were starved, and inspected the expression pattern of each Gal4 driver. Among those candidates, we focused on two Gal4 drivers, VT58471 and VT43147, because a pair of neurons in the dorsal-lateral area of the brain overlapped clearly in the two candidate Gal4 drivers.

### Intersectional approach to define CN neurons at single-cell resolution

To define CN neurons in a single cell resolution, we combined LexA-driven flippase (*LexAop-FLP*) with *UAS-FRT-stop-FRT-effector* to limit the expression of effector in the areas in which Gal4 and LexA lines overlap. This induced the expression of *UAS-FRT-stop-FRT-smGFP*[31], under the control of *CN-Gal4* and another independent *R20F11-LexA* line. *R20F11-LexA* also labels CN neurons and exhibited a defect when crossed to *LexAop-TNT* in the two-choice assay. We also used this approach to induce the expression of UAS-*FRT-stop-FRT-TNT*, under the control of *CN-Gal4* and *R20F11-LexA*, to silence a pair of CN neurons.

### Immunostaining

Immunohistochemistry of the brain, ventral nerve cord (VNC), CC and foregut was conducted as previously described[5]. In brief, the brains were dissected in PBS (1.86 mM NaH2PO4, 8.41 mM Na2HPO4, 175 mM NaCl) and fixed in 4% paraformaldehyde (PFA)/PBS for 30 min at 23 °C. After washing in PBST (PBS + 0.3% Triton X-100, Invitrogen) (3 times, 10 min each), the brains were blocked in 10% normal goat serum (NGS, Jackson Immunoresearch, T-005-000-121) in PBST for 1 h at 23 °C, and then incubated with primary antibodies for 12-48 h at 4 °C. After washing in PBST (3 times, 10 min each), the sample brains were incubated with secondary antibodies for 12-24 h at 4 °C and washed again using PBST (3 times, 10 min each). Primary antibodies that were used: chicken anti-GFP (1:500; Invitrogen, A10262), rabbit anti-GFP (1:500; Invitrogen, A-11122), mouse anti-GFP (1:100; Sigma, G6539, use for synaptobrevin (Syb)-GRASP), mouse anti-nc82 (1:25; Development Studies Hybridoma Bank, DSHB, AB-2314866), rabbit anti-dsRed (1:500; Clontech, 632496), rabbit anti-corazonin (Crz) (1:500; a gift from J. Veenstra, Université de Bordeaux), rabbit anti-sNPF (1:500; a gift from J. Veenstra, Université de Bordeaux, France), rabbit anti-dilp2 (1:500; a gift from E. Hafen, Institute for Molecular Systems Biology, Zürich) and rabbit anti-AKH (1:500; gifts from J. H. Park, University of Tennessee and S. K. Kim, Stanford University) antibodies. Secondary antibodies that were used: Alexa Fluor 633 goat anti-rabbit IgG (1:500; Invitrogen, A-21070), Alexa Fluor 555 goat anti-mouse IgG (1:500; Invitrogen, A-21127), Alexa Fluor 555 goat anti-rabbit IgG (1:500; Invitrogen, A27039), Alexa Fluor 488 goat anti-rabbit IgG (1:500; Invitrogen, A27034), Alexa Fluor 488 goat anti-mouse IgG (1:500; Invitrogen) and Alexa Fluor 488 goat anti-chicken IgG (1:500; Invitrogen, A28175). We used Vectashield (Vector Labs, H1000) for mounting the samples. All Images were acquired using a Zeiss LSM 800 confocal microscope (Zeiss) with 25× lens at 1,024 × 1,024 resolution. Z-stacked images were constructed using ZEN image analysing software (Carl Zeiss, ZEN 2.3 SP1 FP1, v.14.0.12.201). Quantifications and statistical analyses were conducted using ImageJ and GraphPad Prism 8.1.1, respectively.

### Ex vivo GCaMP and Arclight imaging

Ex vivo calcium imaging using GCaMP6s was performed as described[5]. Adult male fly brains or CC cells in the foregut were dissected with AHL (108 mM NaCl, 8.2 mM $MgCl_2$, 4 mM $NaHCO_3$, 1 mM $NaH_2PO_4$, 2 mM $CaCl_2$, 5 mM KCl, 5 mM HEPES, appropriate amount of sucrose to balance osmolarity, the pH was adjusted to 7.3 with 1 M NaOH) and were immobilized by tissue holder (Warner Instruments) on a Sylgard-based perfusion chamber. Sugars (D-glucose, D-trehalose, D-fructose and sucrose), 2DG, phlorizin, pyruvate and L-essential amino acids (EAAs) (1× L-(10)-EAAs: 0.6 mM L-Arg, 0.2 mM L-His, 0.4 mM L-Ile, 0.4 mM L-Leu, 0.4 mM L-Lys, 0.1 mM L-Met, 0.2 mM L-Phe, 0.4 mM L-Thr, 0.05 mM L-Trp and 0.4 mM L-Val) were purchased from Sigma. TTX, glibenclamide and nimodipine were purchased from Tocris. We mixed 20 mM D-glucose containing AHL with 0.5 μM TTX, 20 mM 2DG, 1 mM phlorizin or 5 μM nimodipine. The concentrations of TTX, glibenclamide, 2DG, phlorizin, nimodipine, sucrose, pyruvate and L-essential amino acids (L-(10)-EAAs) that were used in the experiments were determined based on the pilot experiments. Ex vivo Arclight imaging was performed as previously described with a modification[32]. Each brain was recorded for 200–500 frames in total (512 × 512 pixels; each frame, 5 s). After imaging, the condition of the cells was checked by treating 80 mM KCl contained AHL solution. Exchange of different solutions was automatically controlled by a ValveBank controller (AutoMate Scientific). Changes in fluorescent intensity were recorded using a Prairie two-photon microscope and Prairie view software v.4.3.2.18 (Prairie Technologies Inc.) with a 40× water immersion lens (Olympus). All Image analyses were conducted using ImageJ. A region of interest (ROI) was centred using StackReg plugin in ImageJ.

Peak amplitude (% $\Delta F/F$) was obtained by subtracting the amplitude of pre-stimulation baseline (average of 30 frames, 1 frame = 5 s) from the stimulation-evoked peak amplitude. Oscillation number refers to the total number of calcium oscillations during stimulation. Duration is the length of calcium response during stimulation. Oscillation frequency is a ratio of oscillation numbers/duration. Max $\Delta F/F$ (%) = $((F_{max} - F_0)/F_0) \times 100$; $F_{max}$, maximum fluorescence observed during stimulation; $F_0$, average fluorescence of 30 baseline slides. Min $\Delta F/F$

(%) = $((F_{min} - F_0)/F_0) \times 100$; $F_{min}$, minimum fluorescence observed during stimulation. Max $\Delta F/F$ (%) was used when we calculated the increasing amount of fluorescent signal of cells during stimulation, and minimum $\Delta F/F$ (%) was used when we calculated the decreasing amount of fluorescent signal of cells during stimulation. Min $\Delta F/F$ (%) was used only in Fig. 4d. Statistical analyses were conducted with GraphPad Prism 8.1.1.

### In vivo GCaMP imaging

In vivo calcium imaging using GCaMP6s was modified from the previous study[33]. Adult male flies were collected (1–2 days old) and recovered in standard fly food for two more days. They were then fixed on the Sylgard-based perfusion chamber using Epoxy resin (Devcon) and their head cuticles were removed to make a window in order to conduct in vivo calcium imaging. Other procedures are similar with those of ex vivo calcium imaging.

### CaLexA measurement

Adult male flies carrying *Crz-Gal4* (or *CN-Gal4*) and *UAS-CaLexA* (*UAS-mLexA-VP16-NFAT, LexAop-GFP*) were collected 1–2 days after eclosion and recovered for 2–3 more days in standard fly food at 23 °C. A group of flies were dissected and their brains were fixed with 4% PFA in PBS for 20 min without antibody staining (fed flies). Another group of flies after 24 h of starvation (wet starvation) were dissected and their brains were fixed similarly (starved flies). We refed some of the starved flies with 200 mM D-glucose (with 1% agar) for 24 h and their brains were fixed similarly (refed flies). Native CaLexA-driven GFP signals were captured using confocal microscopy (LSM800, Zeiss, 25X lens). The brains were co-stained with anti-Crz antibody. The intensity of CaLexA-driven GFP signal was normalized with the intensity of anti-Crz staining in CN neurons. The CaLexA-driven GFP signal was measured and normalized using corrected total fluorescence (CTF) method[34], calculated by subtracting the background endogenous fluorescence from the integrated GFP signal of an area of interest. *Z*-stacked images were constructed using ZEN software (Zeiss) and analysed using ImageJ.

### Measurement of nutrient-dependent plasticity

To measure nutrient-dependent plasticity of CN axons and presynaptic terminals, adult male flies carrying *Crz-Gal4* (or *CN-Gal4*) and *UAS-Syt::GFP* were collected 1–2 days after eclosion and recovered for 2–3 more days in standard fly food at 23 °C. A group of flies were dissected and their brains were stained using anti-GFP antibody (fed flies). Another group of flies were dissected after 24 h of starvation (wet starvation) and their brains were stained similarly (starved flies). We refed some of the starved flies with 200 mM D-glucose (with 1% agar) for 24 h and their brains were stained similarly (refed flies). We measured and quantified the number of Syt-GFP$^+$ puncta and the length of axons as described in previous study[35]. The intensity of Syt-GFP signal was measured and normalized using corrected total fluorescence (CTF) method[34], calculated by subtracting the background endogenous fluorescence from the integrated GFP signal in an area of interest. Images were captured by Zeiss LSM 800 confocal microscopy (Zeiss) and were analysed using ImageJ. Statistical analyses were conducted with GraphPad Prism 8.1.1.

### GRASP analysis

In Fig. 3c, experimental flies carrying *CN-Gal4, UAS-Neurexin (Nrx)::sGFP1-10, R20F11-LexA* and *LexAop-CD4::sGFP11* and control flies carrying *CN-Gal4, UAS-Neurexin (Nrx)::sGFP1-10* and *LexAop-CD4::sGFP11* were used to examine the physical connectivity between CN neurons and IPCs. We expressed one component of the split GFP with the presynaptic marker neurexin in CN neurons and another component in IPCs. In Fig. 4c, experimental flies carrying *R20F11-LexA, LexAop-Syb::GFP1-10, AKH-Gal4* and *UAS-CD4::GFP11* and control flies carrying *R20F11-LexA, LexAop-Syb::GFP1-10* and *UAS-CD4::GFP11* were used to examine the physical connectivity between CN neurons and AKH-producing cells. We expressed one component of the split-GFP with the presynaptic marker, synaptobrevin, in the CN neurons and another component in the AKH-producing cells. To measure native GRASP-induced GFP signals, adult male flies were collected 3-5 days after eclosion and their brains were dissected in ice-cold PBS and fixed in 4% PFA/PBS for 20 min at 23 °C and washed in PBST (3 times, 10 min each). Then the brains were mounted in Vectashield medium (Vector Labs, H1000) without antibody staining. To measure the synaptobrevin-GRASP[23]-induced GFP signals, adult male flies were collected 3-5 days after eclosion and their brains were dissected in PBS and fixed in 4% PFA/PBS for 30 min at 23 °C. These brains were then blocked using 10% NGS/PBST (1 h at 23 °C), and washed in PBST (3 times, 10 min each) before subsequent immunolabelling with mouse anti-GFP antibody to probe GFP signals as previously described[23]. The brains were mounted in Vectashield medium (Vector Labs, H1000) after completing staining processes.

### Functional connectivity using $P2X_2$ system

Experimental flies carrying *dilp2-LexA*, *LexAop-GCaMP6s*, *CN-Gal4* and *UAS-P2X$_2$* and control flies carrying *dilp2-LexA*, *LexAop-GCaMP6s* and *UAS-P2X$_2$* were used to examine the functional connectivity between CN neurons and IPCs. Experimental flies harbouring *R20F11-LexA, LexAop-P2X$_2$, AKH-Gal4* and *UAS-GCaMP6s* and control flies bearing *LexAop-P2X$_2$, AKH-Gal4* and *UAS-GCaMP6s*, and experimental flies carrying *R20F11-LexA*, *LexAop-P2X$_2$*, *AKH-Gal4* and *UAS-Arclight* and control flies harbouring *LexAop-P2X$_2$*, *AKH-Gal4* and *UAS-Arclight* were used to examine the functional connectivity between CN neurons and AKH-producing cells. To clarify the nutrient dependent changes in the synapses between CN neurons and IPCs, and between CN neurons and AKH-producing cells, we used experimental flies fed with normal food, 24-h-starved (wet starvation) experimental flies, or control flies fed with normal food. We applied 2.5 mM ATP (Sigma, A26209) for 50 s (10 slides, 5 s in each slide) in AHL to excite $P2X_2$ receptor expressing cells. Changes in fluorescent intensity were recorded using a Prairie two-photon microscope with a 40× water immersion lens (Olympus). Image analyses were conducted using ImageJ. A region of interest (ROI) was centred using StackReg plugin in ImageJ. Plotting graphs and statistical analyses were conducted with GraphPad Prism 8.1.1.

### Dilp2 and AKH secretion assay

To measure dilp2 secretion/retention, adult male flies (1–3 days old) were collected under $CO_2$ anaesthesia and allowed to recover for 2–3 more days in standard fly food. A group of flies were starved with water (wet starvation) for 24 h. Their brains were dissected and stained using anti-dilp2 antibody. Some of the starved flies were refed with 200 mM D-glucose (with 1% agar) for 24 h, and their brains were dissected and stained using anti-dilp2 antibody. All the processes were conducted at 23 °C. To measure AKH secretion and retention, adult male flies (1–3 days old) were collected under $CO_2$ anaesthesia and allowed to recover for 2–3 more days in standard fly food. Their brains were dissected and stained using anti-AKH antibody as described. The analysis of Z-stacked images was performed using ZEN software (Zeiss) and ImageJ.

### Quantification of dilp2 and AKH by PRM mass spectrometry

Haemolymph of 500–600 flies that were fed with standard fly food was collected and pooled as previously described[36]. Dilp2 and AKH were extracted from the control (*UAS-Kir2.1/+)* and experimental (*UAS-Kir2.1/CN-Gal4*, CN silenced) samples using two volumes of cold ethanol following as a previously described method with modification[37]; 100 µl of haemolymph was mixed with 200 µl of ice cold ethanol (Fisher Scientific, 90% ethyl alcohol (v/v) with 5% isopropanol and 5% methanol) and incubated at −20 °C for 30 min, then centrifuged at 15,000*g* for 20 min. Supernatants were carefully removed and dried to a small droplet by vacuum centrifugation. Extracts were solubilized in 100 µl of 100 mM ammonium bicarbonate and disulphide bonds reduced with 500 mM dithiothreitol to a final concentration of 10 mM DTT for 1 h at 37 °C. Cysteines were alkylated with iodoacetamide (200 mM) to a final

concentration of 40 mM at 37 °C for 1 h in the dark. Excess iodoacetamide was quenched by adding DTT to a final concentration of 40 mM with incubation for 30 min, in the dark, at room temperature. Samples were each digested with 800 ng to 1 µg of trypsin (Trypsin Gold, Mass Spectrometry Grade, Promega) overnight, and after acidification with 10% formic acid (final concentration of 0.5–1% formic acid), resulting peptides were desalted using hand packed reversed phase Empore C18 Extraction Disks (3M) using a previously described method[38]. Desalted peptides were concentrated to a small droplet by vacuum centrifugation and reconstituted in 10 µl 0.1% formic acid in water. 10% of the peptide material was used for data-dependent-acquisition (DDA) and 50% used for targeted parallel reaction monitoring (PRM) liquid chromatography followed by tandem mass spectrometry (LC-MS/MS). A Q Exactive HF mass spectrometer was coupled directly to an EASY-nLC 1000 (Thermo Fisher Scientific) equipped with a self-packed 75 µm × 20 cm reverse phase column (ReproSil-Pur C18, 3 µM, Dr. Maisch) for peptide separation. Analytical column temperature was maintained at 50 °C by a column oven (Sonation). Peptides were eluted with a 3–40% acetonitrile gradient over 110 min at a flow rate of 250 nl $min^{-1}$. For DDA, the mass spectrometer was operated in DDA mode with survey scans acquired at a resolution of 120,000 (at $m/z$ 200) over a scan range of 300–1750 $m/z$. Up to 15 most abundant precursors from the survey scan were selected with an isolation window of 1.6 Th and fragmented by higher-energy collisional dissociation with normalized collision energy (NCE) of 27. The maximum injection time for the survey and MS/MS scans was 60 ms and the ion target value (AGC) for both scan modes was set to 3e6. For PRM analysis, 1 full MS scan was acquired at 60,000 resolution followed by MS/MS of 9 target precursor $m/z$ loaded as the inclusion list. Each PRM target peptide was analysed at resolution 15,000 with isolation window of 1.4 Th. The ion target value was set to 5e5.

**Mass spectrometry data processing**

RAW files generated from the DDA experiments were analysed by MaxQuant proteomics software[39] (v.1.5.7.0) using a *Drosophila* fasta (19,694 entries) database. Files from PRM experiments were analysed with Thermo Scientific Xcalibur (v.4.1.31.9) software. Layouts containing target precursor and fragment masses were created with a mass accuracy set to 5 ppm. At specific events when MS1 and MS/MS fragments were aligned, the intensity of each specific precursor and fragment were extracted and noted. Additional confirmation of the analysis was carried out using Skyline (v.4.10.18169) Proteomics software[40]. Ion intensity extraction and alignment of spectra were visually inspected and confirmed in each case. Processed intensities of dilp2 B chain, AKH and those fragment ions were normalized by the estimated total protein level (top 3 precursor intensity of top 99 proteins) in the haemolymph. Normalization of precursor peptide and fragment intensities were done by analysing the DDA spectra from each set of replicate control and experimental samples by MaxQuant label-free quantitative analysis; intensities of top three peptides from proteins found in both samples were summed and compared (see Supplementary Table 3). The nomenclature for peptide fragmentation during mass spectrometry was described in previous studies[41,42].

Note: We also performed several experiments using methanol to extract peptides from the haemolymph[43]. The dilp2 and AKH quantitation results from those PRM experiments were consistent with results obtained using the ethanol extracted materials presented in this manuscript, but contained additional contaminants that interfered with liquid chromatography (data not shown).

**Dot blot assay**

Haemolymph of 60–80 flies that were fed with standard fly food was collected as previously described[36]. The total protein in the haemolymph was measured by Bradford (Bio-Rad) and the concentration was adjusted to 1 mg $ml^{-1}$ in PBS[44]. 10 µl of haemolymph was dropped on 0.2 µm nitrocellulose membrane (GE Healthcare) and left at room temperature until to be dried. The membrane was then boiled in PBS for 3 min and subsequently fixed in 4% PFA in PBS for another 20 min. The membrane was blocked with 3% bovine serum albumin (BSA) in PBS for 1 h at room temperature (21 °C) and then incubated with mouse anti-HA (1:500; Covance, 901501) antibody or purified rabbit anti-AKH antibody (1:500; a gift from S. K. Kim, Stanford University) in 3% BSA in PBS at 4 °C for overnight followed by incubation with horseradhish peroxidase-conjugated secondary antibodies in 3% BSA in PBS for 1 h at room temperature (21 °C). $P_{dilp2}$-*dilp2-HA*[36] line was used for detecting dilp2 in the haemolymph. The membrane was developed using Chemidoc M (Bio-Rad). The intensities of the black dots were considered as the amounts of dilp2 and AKH in flies. Ponceau staining was used as loading control. The quantification and analysis of dot blot results were conducted using Fiji software and GraphPad Prism 8.1.1, respectively. For gel source data, see Supplementary Fig. 1.

**Measurement of intracellular sNPF and Crz levels in CN neurons and identification of CN axons and dendrites**

Intracellular sNPF and Crz levels after incubating the brains with sugars were measured as previously described[5]. The brains of 18-h-starved flies in which *UAS-mCD8::GFP* was expressed in Crz-expressing neurons (*Crz-Gal4* > *UAS-mCD8::GFP*) were dissected in cold AHL (sugar free), incubated in 80 mM sucrose, 80 mM D-glucose (D-Glc), 80 mM D-glucose mixed with 0.5 µM TTX (D-Glc + TTX) or 80 mM L-glucose (L-Glc) contained AHL for 30 min and then fixed with 4% PFA/PBS and stained with anti-sNPF or anti-Crz antibodies as described earlier. All Images were acquired using a Zeiss LSM 800 confocal microscope (Zeiss) with 25× lens at 1,024 × 1,024 resolution. *Z*-stacked images were constructed using ZEN image analysing software (Zeiss). Quantifications and statistical analyses were conducted using ImageJ and GraphPad Prism 8.1.1, respectively. To identify the axons and dendrites of CN neurons, we expressed *UAS-Syt::eGFP* under the control of CN-Gal4 line to visualize the axons and *UAS-DenMark*[45] to visualize the dendrites. We stained the brain and gut tissue with anti-GFP antibody to detect the signals driven by *UAS-Syt::eGFP* and anti-dsRed antibody to detect the signals driven by *UAS-DenMark* DenMark (Dendritic Marker, ΔICAM5–Cherry, a hybrid protein of ICAM5 (also known as telencephalin) and mCherry) is a specially designed protein to label dendrites. We used a dsRed antibody to visualize mCherry.

**Classification of the IPCs into three subpopulations**

When max $\Delta F/F$ (%) of a cell is higher than 50% of average max $\Delta F/F$ (%) in control brains, we counted the cell as a 'stimulated' cell. We classified the IPCs into stimulated or unresponsive cells after treating D-glucose or $K_{ATP}$ channel blocker to the brains of flies in which CN neurons had been silenced or in which *sNPF* receptor in IPCs had been inactivated, or control flies. Using this approach, we categorized IPCs into three subpopulations. In Fig. 3h, 74.42% of IPCs (32/43 cells) in control flies carrying *R20F11-LexA*, *dilp2-Gal4* and *UAS-GCaMP6s* responded to D-glucose, whereas 25.58% of the IPCs (11/43 cells) failed to respond to D-glucose. 20.69% of IPCs (12/58 cells) responded to D-glucose in experimental flies carrying *R20F11-LexA*, *LexAop-TNT*, *dilp2-Gal4* and *UAS-GCaMP6s*, whereas 53.73% of the IPCs failed to respond to glucose. In Extended Data Fig. 8d, 71.43% of IPCs (20/28 cells) in control flies carrying *R20F11-LexA*, *dilp2-Gal4* and *UAS-GCaMP6s* responded to glibenclamide, whereas 28.57% of the IPCs (8/28 cells) failed to respond to glibenclamide. 23.08% of IPCs (9/39 cells) in experimental flies carrying *R20F11-LexA*, *LexAop-TNT*, *dilp2-Gal4* and *UAS-GCaMP6s* responded to glibenclamide, whereas 48.35% of the IPCs failed to respond to glibenclamide. In Extended Data Fig. 8g, 75.68% of IPCs (28/37 cells) in control flies carrying *dilp2-Gal4* and *UAS-GCaMP6s* responded to D-glucose, whereas 24.32% of the IPCs (9/37 cells) failed to respond. 27.27% of IPCs (9/33 cells) responded to D-glucose in experimental flies carrying *dilp2-Gal4*, *UAS-sNPFR-DN* and *UAS-GCaMP6s*, whereas 48.41% of the IPCs failed to respond.

### sNPF neuropeptide treatment to the IPCs and CC

sNPF neuropeptide was administrated to IPCs and CC using a modified method from a previous study[26]. sNPF was purchased from Thermo Fisher Scientific (sequence: AQRSPSLRLRF, purity >95%, unmodified). We made 50 mM stock solution of sNPF in DMSO (Sigma, D2650) and keep it at -80 °C. Pertussis toxin (Tocris, 3097), U73122 (Tocris, 1268) or U73343 (Tocris, 4133) was mixed with D-glucose or sNPF. These mixtures were applied to IPCs in the brain or AKH-producing cells in CC. We mixed 80 μM sNPF or 20 mM D-glucose contained AHL with 1 ng $\mu l^{-1}$ PTX, 1 μM U73122 or 1 μM U73343 (a non-functional enantiomer of U73122). The concentrations of PTX, U73122 and U73343 that were used in the experiments were determined based on the pilot experiments.

### Haemolymph glycaemia measurement

Haemolymph glucose and trehalose levels were measured as previously described[5]. In brief, 30–40 flies starved with water for 24 h (starved) or fed with standard fly food (fed) were decapitated and their haemolymph was collected with a capillary pipette (0.25 μl Microcaps, Drummond). 0.25 μl of the haemolymph was mixed with 50 μl of Glucose (HK) Assay Kit (Sigma, GAHK20) and incubated for 20 min at 23 °C, and then measured the absorbance at 340 nm using a Nanodrop spectrophotometer (Thermo Scientific). To measure trehalose concentrations in the haemolymph, pig kidney trehalase (1:500, Sigma, T8778) was added to the mixture of Glucose Assay Kit and incubated for 16–20 h at 37 °C, and then measured the absorbance at 340 nm. Standard curves (D-glucose and D-trehalose) were generated for each trial. Plotting graphs and statistical analyses were conducted with GraphPad Prism 8.1.1.

### Statistics and reproducibility

All statistical analyses were conducted with GraphPad Prism 8.1.1 (provided by NYU School of Medicine). To compare two normally distributed groups, unpaired two-tailed *t*-tests were used. For multiple comparisons between normally distributed groups, one-way ANOVAs followed by Tukey's post hoc test were used (see Supplementary Table 1). Without an asterisk means non-significant ($P > 0.05$). No statistical methods were used to predetermine sample size. The statistical analyses presented in this manuscript were assisted by a statistics expert (X. Li) in NYU. For the two-choice behavioural data (Figs. 1b, e, 5a, b, Extended Data Figs. 1a, d, 2c, 4a–c), each data point (dot) represents a biological replicate bearing 30–40 male flies for a trial. For ex vivo imaging data using GCaMP6s or Arclight (Figs. 2c, d, 3d, g, 4d, e,5d, f, g, Extended Data Figs. 3b–e, 3g–j, 4e–h, 6b, 8c, 10b), each data point (dot) represents a biological replicate of maximum or minimum amplitudes of $\Delta F/F$ (%) during the experimental period. In Extended Data Fig. 6c–f, each average trace contains at least three biological replicates. For confocal fluorescent images of the brains or CCs stained with antibodies (Figs. 2e–g, 3a–c, e, 4a–c, f,5c, e, Extended Data Figs. 4i, j, 5, 9e), we present the exact number of biological replicates in Supplementary Table 1. Moreover, the representative images (Figs. 1a, c, d,2a, 3c, 4c, Extended Data Figs 1c, 2a, b, 9a, b) were independently replicated more than five times. For the mass spectrophotometry and dot blot assay (Figs. 3f, 4g, Extended Data Fig. 7g–j), each data point (dot) or a trial represents a biological replicate bearing 500–600 flies for mass spectrophotometry and 60–80 flies for dot blot assay. For the two-choice behavioural and ex vivo imaging experiments, we repeated more than three times based on our pilot experiments. For the behavioural experiments, we did not include two or three independent pilot experiments because the experimental setups were not clearly controlled at those times. However, they had qualitatively similar results; yet, the standard deviations were substantially high.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

Raw mass spectrometry files have been deposited in the MassIVE database; with MassIVE accession ID: MSV000083796. All other raw data are available from the corresponding author on reasonable request.


31. Nern, A., Pfeiffer, B. D. & Rubin, G. M. Optimized tools for multicolor stochastic labeling reveal diverse stereotyped cell arrangements in the fly visual system. *Proc. Natl Acad. Sci. USA* **112**, E2967–E2976 (2015).
32. Jourjine, N., Mullaney, B. C., Mann, K. & Scott, K. Coupled sensing of hunger and thirst signals balances sugar and water consumption. *Cell* **166**, 855–866 (2016).
33. Ai, M. et al. Acid sensing by the *Drosophila* olfactory system. *Nature* **468**, 691–695 (2010).
34. Burgess, A. et al. Loss of human Greatwall results in G2 arrest and multiple mitotic defects due to deregulation of the cyclin B–Cdc2/PP2A balance. *Proc. Natl Acad. Sci. USA* **107**, 12564–12569 (2010).
35. Liu, Q. et al. Branch-specific plasticity of a bifunctional dopamine circuit encodes protein hunger. *Science* **356**, 534–539 (2017). https://doi.org/10.1126/science.aal3245.
36. Park, S. et al. A genetic strategy to measure circulating *Drosophila* insulin reveals genes regulating insulin production and secretion. *PLoS Genet*. **10**, e1004555 (2014).
37. Chen, W., Hwang, Y. Y., Gleaton, J. W., Titus, J. K. & Hamlin, N. J. Optimization of a peptide extraction and LC–MS protocol for quantitative analysis of antimicrobial peptides. *Future Sci*. **5**, FSO348 (2018).
38. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc*. **2**, 1896–1906 (2007).
39. Cox, J. et al. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res*. **10**, 1794–1805 (2011).
40. MacLean, B. et al. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **26**, 966–968 (2010).
41. Johnson, R. S., Martin, S. A., Biemann, K., Stults, J. T. & Watson, J. T. Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine. *Anal. Chem*. **59**, 2621–2625 (1987).
42. Roepstorff, P. & Fohlman, J. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biol. Mass Spectrom*. **11**, 601 (1984).
43. Gáliková, M. et al. Energy homeostasis control in *Drosophila* adipokinetic hormone mutants. *Genetics* **201**, 665–683 (2015).
44. Saadipour, K. et al. The transmembrane domain of the p75 neurotrophin receptor stimulates phosphorylation of the TrkB tyrosine kinase receptor. *J. Biol. Chem*. **292**, 16594–16604 (2017).
45. Nicolaï, L. J. J. et al. Genetically encoded dendritic marker sheds light on neuronal connectivity in *Drosophila*. *Proc. Natl Acad. Sci. USA* **107**, 20553–20558 (2010).



**Acknowledgements** We thank M. Kanai and J.-Y. Park for providing expertise on the two-choice behaviour assay and physiology. We also thank S. K. Kim, J. Veenstra, J. H. Park and E. Hafen for providing antibodies used in this study; B. Levin, G. Schwartz, J. W. Sohn, H. Kim and the Suh laboratory members for stimulating discussion. This work is supported by NIH R01 grants (R01DK116294, R01DK106636 to G.S.B.S. and NINDS P30NS050276 to T.A.N.), KAIST Chancellor's fund and a grant from Samsung Science & Technology Foundation (G.S.B.S).


**Author contributions** Y.O. performed nearly all immunohistochemistry, calcium imaging, behaviour testing, haemolymph glycaemia measurement, statistical analyses and figure design. J.S.-Y.L., H.J.M. and B.G. carried out the Gal4 screen using the two-choice assay. H.E.B. and T.A.N. conducted mass spectrometry analyses to measure dilp2 and AKH levels in haemolymph. K.S. performed dot blot assays to measure dilp2 and AKH levels in haemolymph. J.G.W. performed feeding assays and assisted in measuring haemolymph glycaemia. F.A. assisted with the CaLexA experiment. G.S.B.S. supervised the project and provided intellectual support. Y.O. and G.S.B.S. wrote the manuscript with inputs from other authors.

**Competing interests** The authors declare no competing interests.

**Additional information**

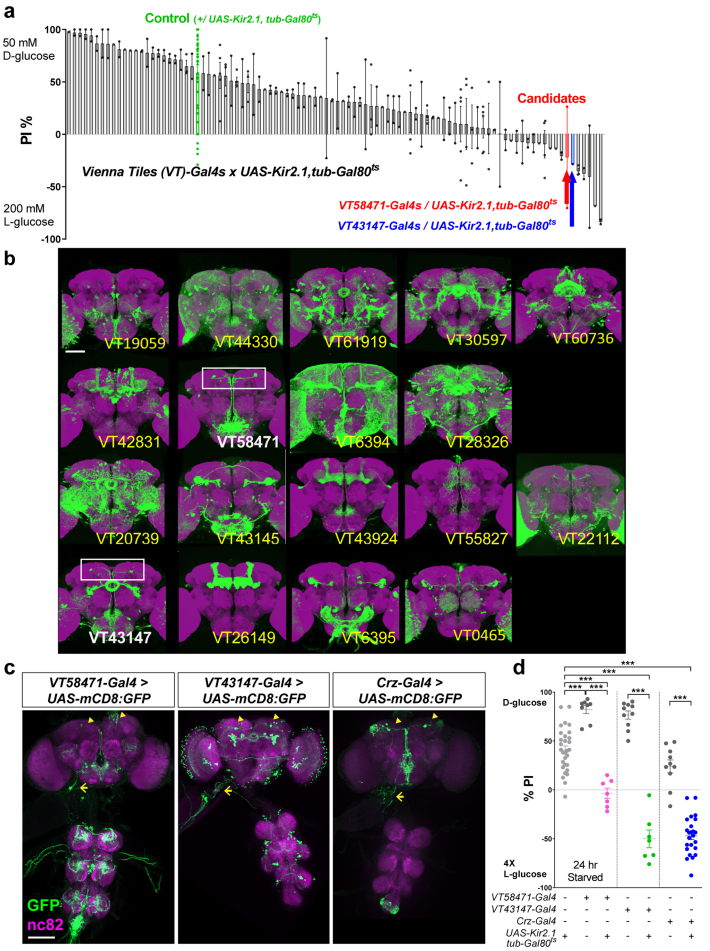

**Supplementary information** is available for this paper at https://doi.org/10.1038/s41586-019-1675-4.



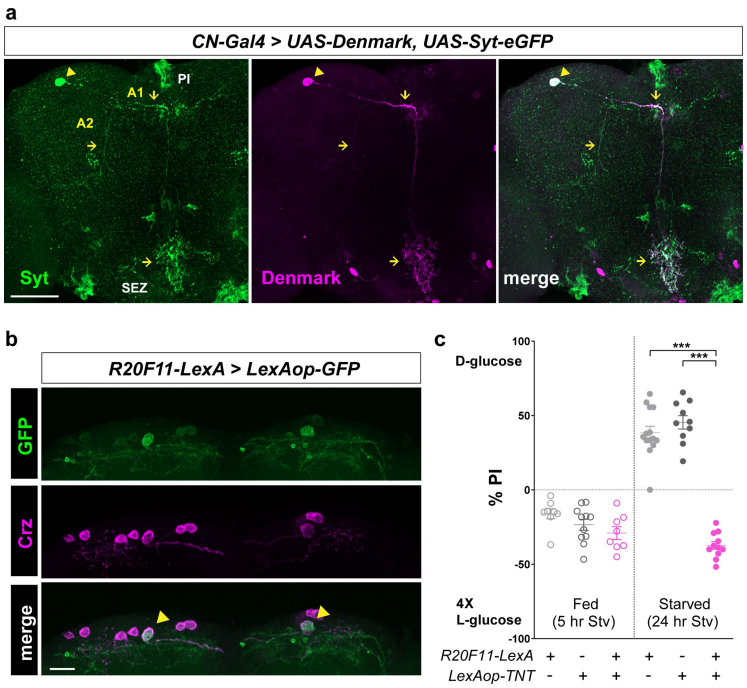
**Correspondence and requests for materials** should be addressed to G.S.B.S.



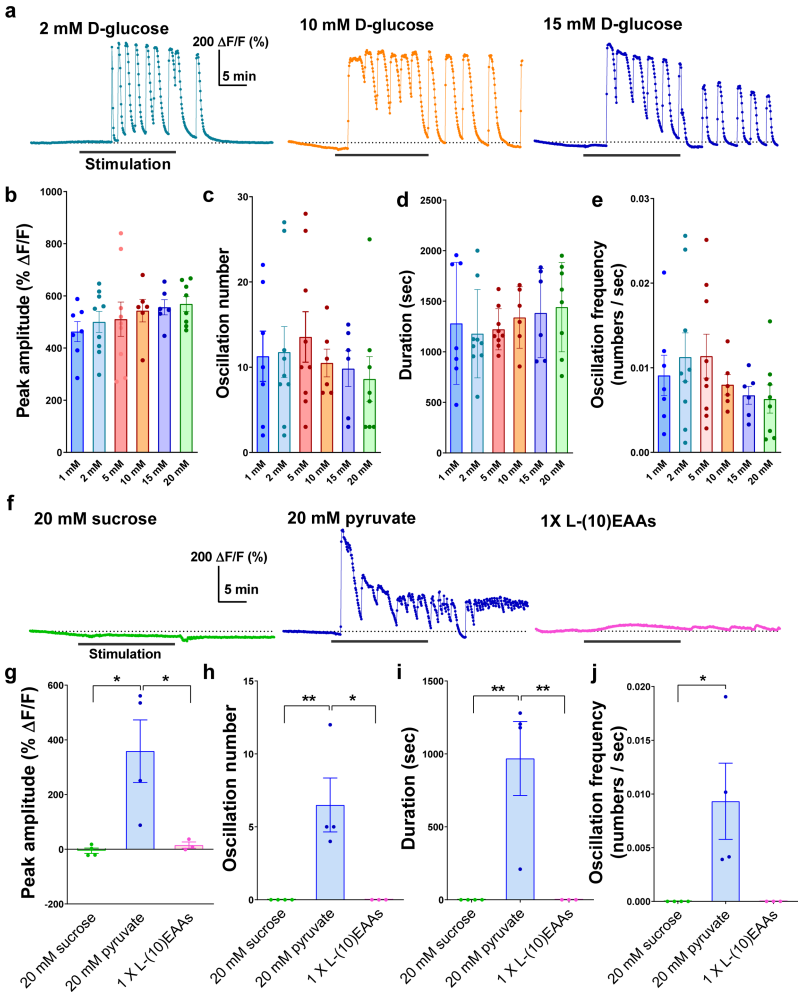
**Peer review information** *Nature* thanks Mark L. Evans and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



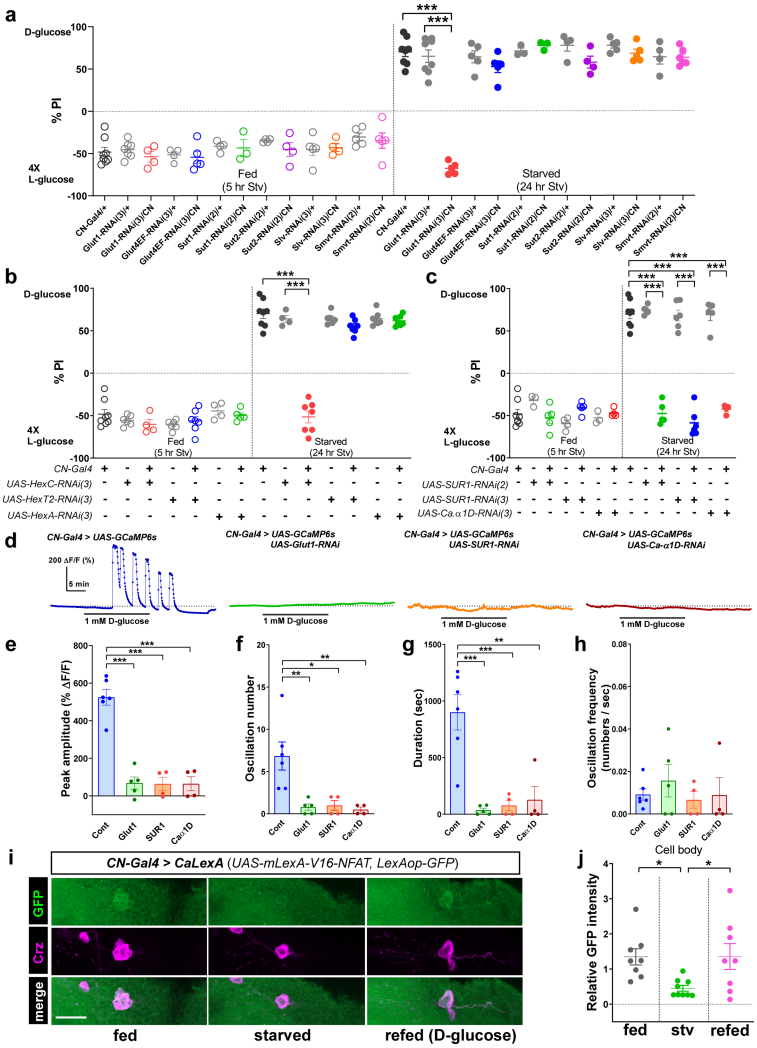

**Extended Data Fig. 1 | Identification of neurons that are required for the starvation-induced nutrient selection. a**, The result of a screen of VT-Gal4 lines when bearing *UAS-Kir2.1, tub-Gal80$^{ts}$* that failed to select D-glucose after 24 h of starvation. **b**, Expression patterns of 18 candidate Gal4 lines isolated from the screen. Scale bar, 100 µm. Brain images from FlyLight Gal4 LexA collection are shown. White boxes denote a pair of the dorsolateral cells. **c**, Expression patterns of *VT58471-Gal4>UAS-mCD8::GFP* line, *VT43147-Gal4>UAS-mCD8::GFP* line and *Crz-Gal4>UAS-mCD8::GFP* line in the brain, the VNC and a part of the foregut, stained with anti-GFP (green) antibody. Neuropil is stained with nc82 (bruchpilot, magenta) antibody. Cell bodies of candidate neurons (arrowheads) and their projections to the foregut (arrows) are shown. Scale bar, 100 µm. Images are *z*-stacked projections. **d**, Inactivation of the candidate neurons by expressing *UAS-Kir2.1, tub-Gal80$^{ts}$* under the control of *VT58471-Gal4, VT43147-Gal4* or *Crz-Gal4* at 30 °C blunts a preference for D-glucose in starved flies. ***$P < 0.001$; one-way ANOVA with Tukey post hoc test. See Supplementary Table 1 for the sample sizes and statistical analyses.



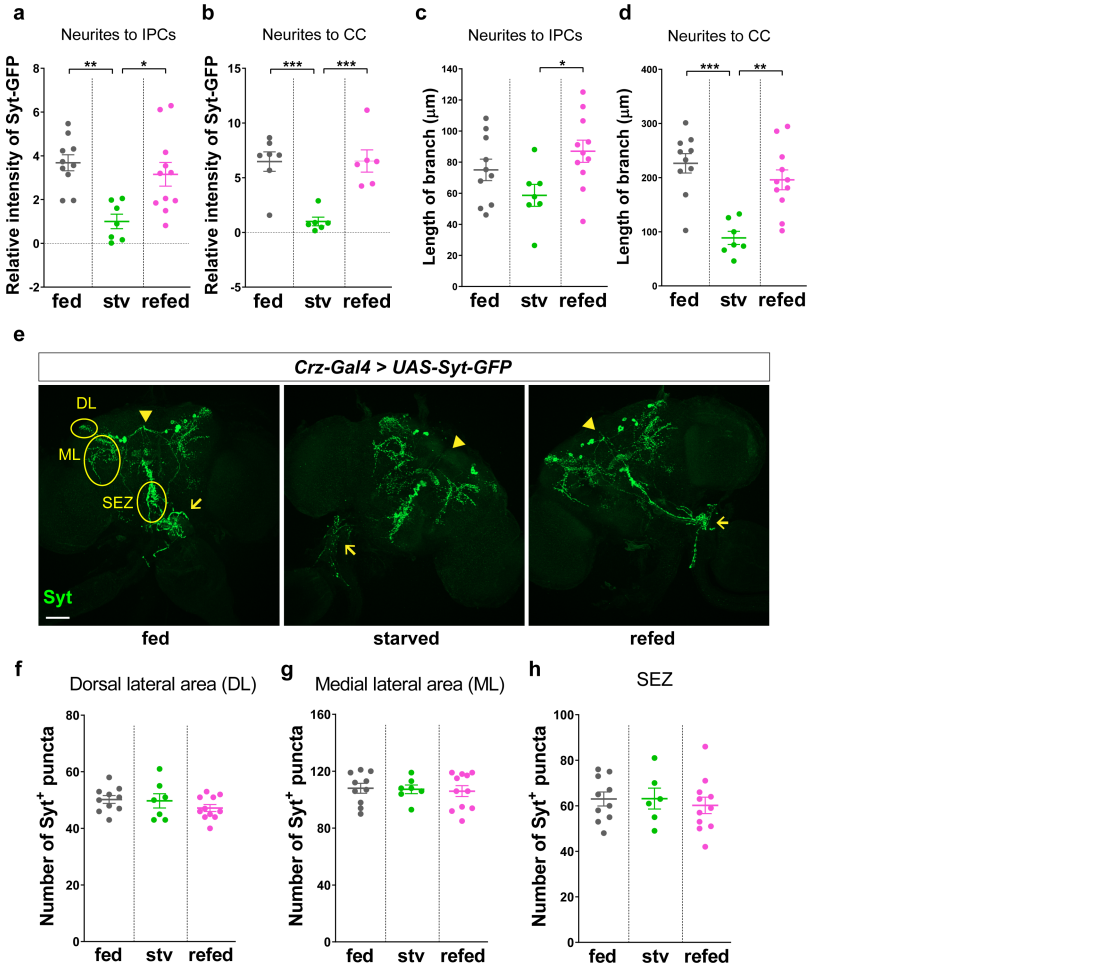

**Extended Data Fig. 2 | Axons and dendrites of CN neurons and a LexA line that labels a pair of CN neurons. a**, Axons and dendrites of CN neurons in the brain visualized by an axonal marker, *UAS-Syt::eGFP*, and a dendritic marker, *UAS-DenMark*, under the control of *CN-Gal4*, stained with anti-GFP (green) and anti-dsRed (DenMark, magenta, see Methods) antibodies. Arrowheads denote CN cell bodies and arrows indicates CN axons (left) and dendrites (middle). **b**, The brain of a fly carrying *R20F11-LexA* and *LexAop-GFP*, stained with anti-GFP (green) antibody. Cell bodies of CN neurons are stained with anti-Crz (magenta) antibody. Arrowheads indicate cell bodies of the dorsolateral CN neurons. Scale bar, 20 µm. **c**, Inactivation of CN neurons by expressing *LexAop-TNT* under the control of *R20F11-LexA* blunts a preference for D-glucose in starved flies. SEZ, subesophageal zone. Scale bar, 50 µm. Z-stacked projections are shown. ***$P < 0.001$; one-way ANOVA with Tukey post hoc test. See Supplementary Table 1 for the sample sizes and statistical analyses.



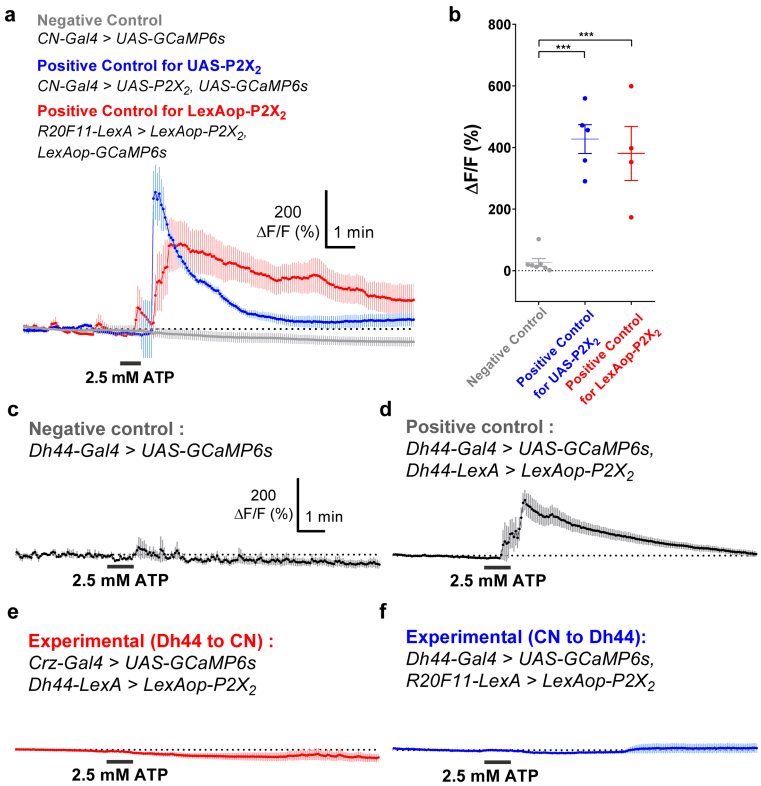

**Extended Data Fig. 3 | The activity of CN neurons is stimulated by glucose or pyruvate, but not by sucrose or a mixture of essential amino acids.** **a**, Representative traces of CN neuronal activity in response to 2 mM, 10 mM and 15 mM D-glucose. **b**–**e**, Quantifications of CN responses to 1 mM, 2 mM, 5 mM, 10 mM, 15 mM and 20 mM D-glucose. When 1 to 5 mM D-glucose were applied, all parameters increased. When higher concentrations (10 to 20 mM D-glucose) were applied, the oscillation number (**c**) and oscillation frequency (**e**) decreased, while the peak amplitude (**b**) and duration (**d**) increased. **f**, Representative traces of CN neuronal activity in response to 20 mM sucrose, 20 mM pyruvate and a mixture of ten EAAs. **g**–**j**, Quantifications of CN neuronal responses to 20 mM sucrose, 20 mM pyruvate and a mixture of 10 EAAs. $^{*}P < 0.05$ and $^{**}P < 0.01$; one-way ANOVA with Tukey post hoc. See Supplementary Table 1 for the sample sizes and statistical analyses.



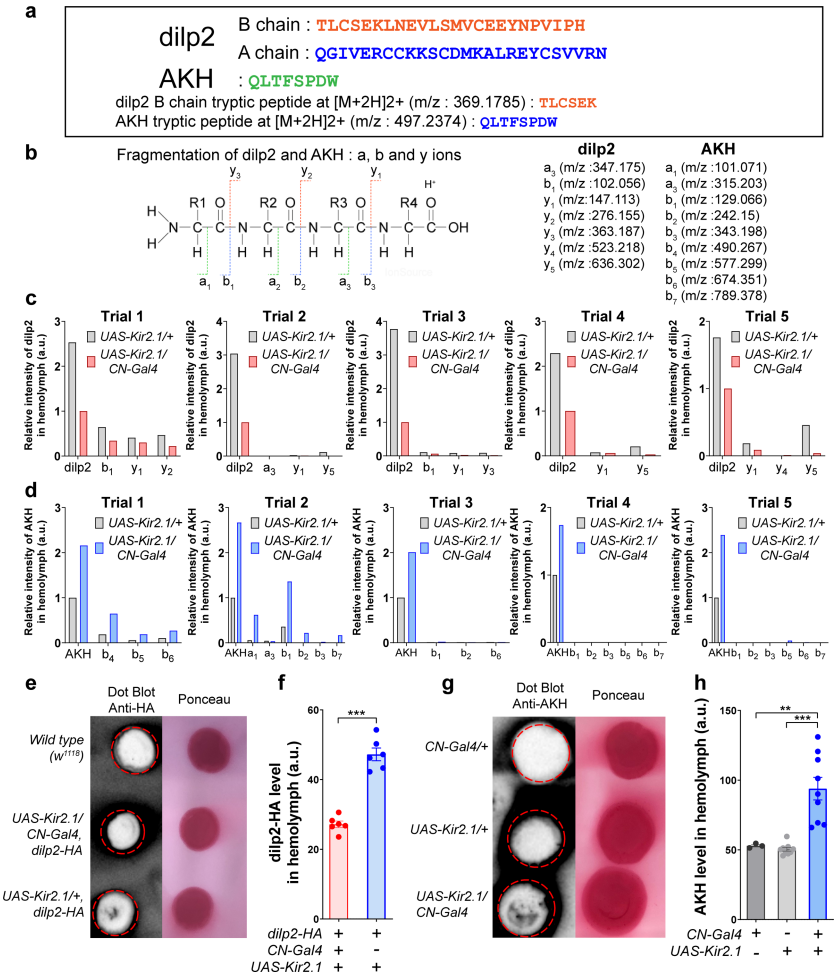

**Extended Data Fig. 4 | *Glut1*, *Hex-C*, $K_{ATP}$ channel and voltage-gated calcium channel are required in CN neurons to respond to glucose, and the activity of CN neuron is controlled by the internal energy state in live animals. a**, Inactivation of *Glut1* by expressing *UAS-Glut1* RNAi under the control of *CN-Gal4* (but not other glucose transporters) blunts a preference for D-glucose in starved flies. **b**, Inactivation of *Hex-C* by expressing *UAS-HexC* RNAi under the control of *CN-Gal4* blunts a preference for D-glucose, but not other hexokinases, in starved flies (*Hex-T2* and *Hex-A*). **c**, Inactivation of $K_{ATP}$ channel or voltage-gated calcium channel by expressing *UAS-SUR1* RNAi or *UAS-Ca-α1D* RNAi by *CN-Gal4* blunts a preference for D-glucose in starved flies. **d**–**h**, Representative traces (**d**) and quantifications (**e**–**h**) of calcium responses to 1 mM glucose by CN neurons of flies in which *Glut1*, *SUR1* or voltage-gated calcium channel subunit (*Ca-α1D*) was knocked down by RNAi, or those of control flies. **i**, **j**, Representative images (**i**) and quantifications of native CaLexA-driven GFP intensity from CN cell bodies (**j**) of fed, starved or refed flies carrying *CN-Gal4* and *UAS-CaLexA* (*UAS-mLexA-VP16-NFAT, LexAop-GFP*). Cell bodies of CN neurons are stained with anti-Crz (magenta) antibody. Scale bar, 20 µm. Images are *z*-stacked projections. $^{*}P < 0.05$, $^{**}P < 0.01$ and $^{***}P < 0.001$; one-way ANOVA with Tukey post hoc test. See Supplementary Table 1 for the sample sizes and statistical analyses.



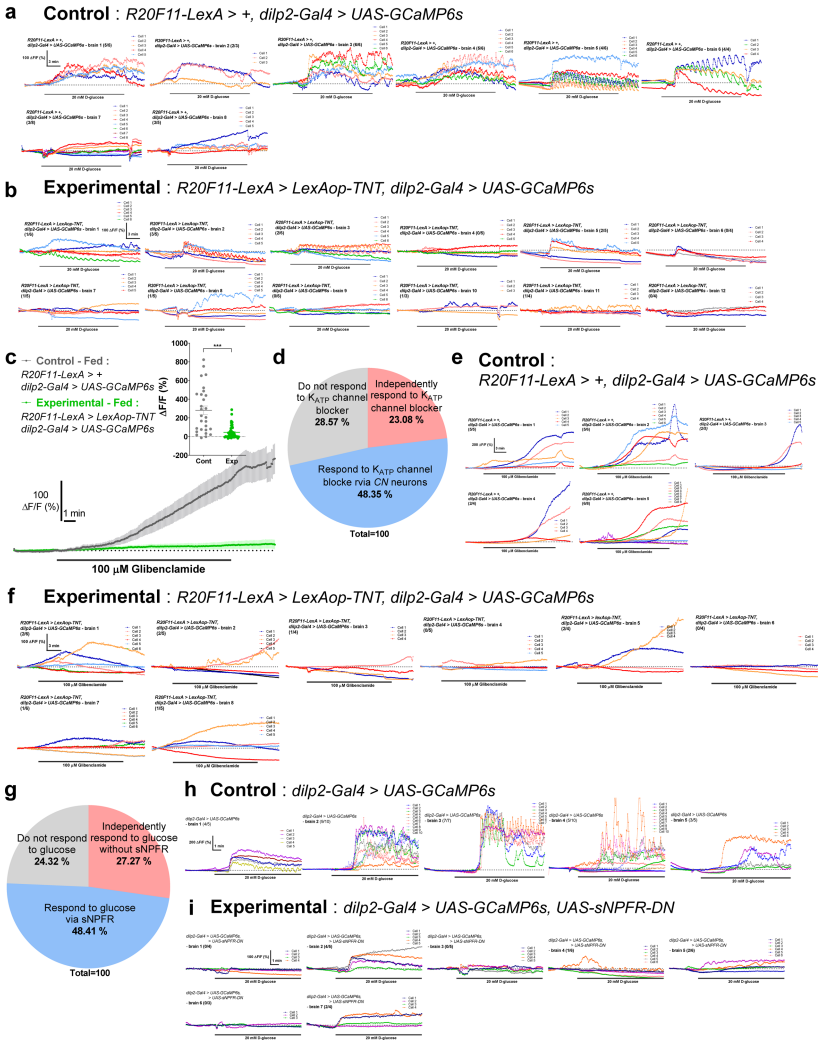

**Extended Data Fig. 5 | Nutrient-dependent plasticity does not occur in the majority of Crz neurons except a pair of the dorsolateral CN neurons.** **a**–**d**, The relative fluorescence intensities of Syt–GFP signals (**a**, **b**) and the branch lengths (**c**, **d**) of axon 1 (A1) (neurite to IPCs, **a**, **c**) and axon 2 (A2) (neurite to CC, **b**, **d**) in fed, starved or refed flies carrying *Crz-Gal4* and *UAS-Syt::GFP*. **e**–**h**, Representative images (**e**) and quantifications (**f**–**h**) of the number of Syt–GFP$^+$ puncta in fed, starved or refed flies carrying *Crz-Gal4* and *UAS-Syt::GFP* in the dorsal lateral area (DL, **f**), the medial lateral area (ML, **g**) and the SEZ. In **h**, Arrowheads and arrows indicate Syt–GFP$^+$ puncta on A1 and A2, respectively. Scale bar, 50 μm. Images are *z*-stacked projections. *$P < 0.05$, **$P < 0.01$ and ***$P < 0.001$; one-way ANOVA with Tukey post hoc test. See Supplementary Table 1 for the sample sizes and statistical analyses.



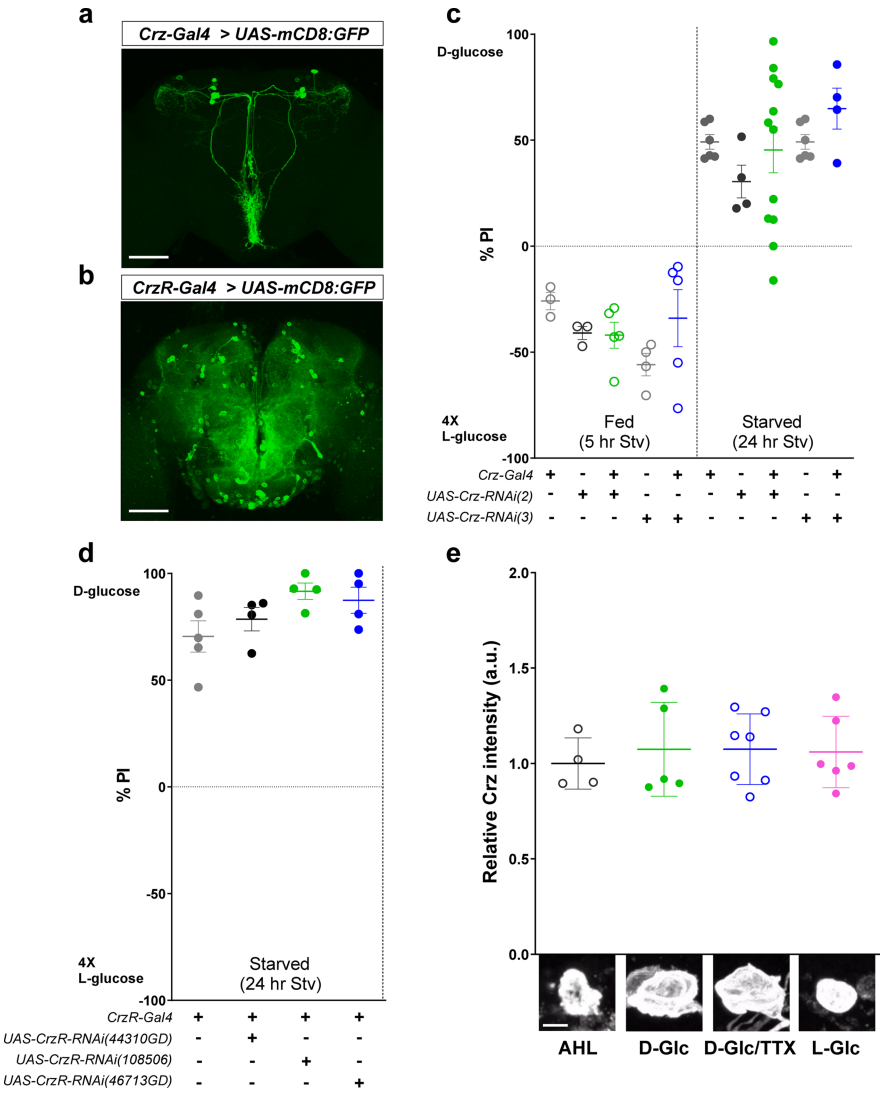

**Extended Data Fig. 6 | CN neurons expressing $P2X_2$ receptor can be readily activated by ATP application, and Dh44 neurons and CN neurons are not functionally coupled. a**, **b**, Average GCaMP traces (**a**) and their Δ*F*/*F* (max) quantifications (**b**) of CN neurons when exposed to 2.5 mM ATP. **c**, Average GCaMP trace of flies carrying *Dh44-Gal4, UAS-GCaMP6s* in response to ATP. **d**, Average GCaMP trace of flies carrying *Dh44-Gal4, UAS-GCaMP6s, Dh44-LexA* and *LexAop-P2X*$_2$ in response to ATP. **e**, Average GCaMP trace of flies carrying *Crz-Gal4, UAS-GCaMP6s, Dh44-LexA* and *LexAop-P2X*$_2$ in response to ATP. **f**, Average GCaMP trace of flies carrying *Dh44-Gal4, UAS-GCaMP6s, R20F11-LexA* and *LexAop-P2X*$_2$ in response to ATP. ***$P < 0.001$; one-way ANOVA with Tukey post hoc test. See Supplementary Table 1 for the sample sizes and statistical analyses.



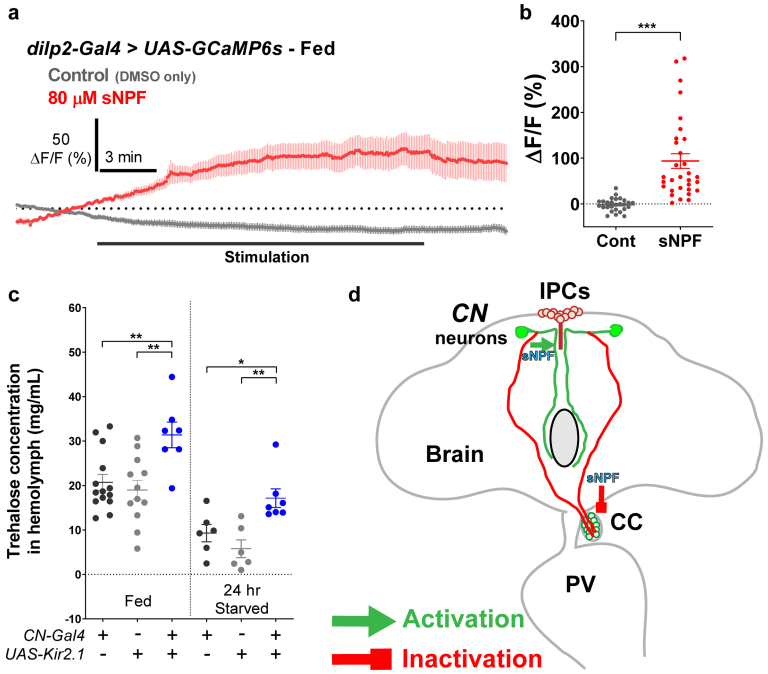

**Extended Data Fig. 7 | The circulating levels of dilp2 and AKH in haemolymph are measured by mass spectrometry and dot blot assay.** **a**, Sequences of dilp2, AKH, tryptic peptide of dilp2 B chain and tryptic peptide of AKH. We detected dilp2 B chain at *m/z* (mass to charge ratio): 369.1785 (TLCSEK, M2H+: 369.1785) and AKH at *m/z*: 497.2374 (QLTFSPDW, M2H+: 497.2374). **b**, Nomenclature and *m/z* values of fragment ions (N-terminal directed 'a' and 'b' ions, as well as, C-terminal directed 'y' ions) which is driven by dilp2 B chain and AKH. **c**, **d**, Relative extracted ion intensities for the dilp2 B chain and its fragment ions in each trial (**c**), and AKH and its fragment ions in each trial (**d**) generated from the haemolymph of fed flies in which CN neurons were inactivated, or those of control flies; see Methods. **e**, **f**, A dot blot (**e**) and its quantification (**f**) show the levels of dilp2 in the haemolymph of wild type ($w^{1118}$), *UAS-Kir2.1/CN-Gal4;dilp2-HA* and *UAS-Kir2.1/+;dilp2-HA* flies, probed with anti-HA antibody to detect dilp2. Because $w^{1118}$ flies do not express dilp2–HA, they were used as a negative control. **g**, **h**, A dot blot (**g**) and its quantification (**h**) show the levels of AKH in the haemolymph of *CN-Gal4/+*, *UAS-Kir2.1/+* and *UAS-Kir2.1/CN-Gal4* flies, probed with anti-AKH antibody. The intensity of black dots in the red dashed circle represents the quantity of dilp2 or AKH that was later normalized to Ponceau staining. For gel source data, see Supplementary Fig. 1. $^{**}P < 0.01$ and $^{***}P < 0.001$; unpaired two-tailed *t*-test (**f**) and one-way ANOVA with Tukey post hoc test (**h**). See Supplementary Table 1 for the sample sizes and statistical analyses.




**Extended Data Fig. 8 | Calcium responses to D-glucose or $K_{ATP}$ channel blocker by IPCs of flies in which CN neurons were silenced or sNPF receptor was rendered nonfunctional. a**, **b**, Individual traces of IPCs in control flies carrying *R20F11-LexA, dilp2-Gal4* and *UAS-GCaMP6s* (**a**) and in experimental flies carrying *R20F11-LexA, LexAop-TNT, dilp2-Gal4* and *UAS-GCaMP6s* (**b**) responding to D-glucose. **c**, Average GCaMP traces and $\Delta F/F$ (max) quantifications from IPCs of fed flies in which CN neurons were inactivated by TNT in response to $K_{ATP}$ channel blocker glibenclamide, or those of control flies. **d**, IPCs partition into three subpopulations depending on their response to glibenclamide; see Methods. **e**–**f**, Individual traces of IPCs in control flies (**e**) and in flies in which CN neurons were inactivated by TNT (**f**) responding to glibenclamide. Because a saturating concentration of D-glucose (20 mM) was used to quantify the populations of IPCs, we used 100 µM glibenclamide, a saturating concentration for glibenclamide according to our control experiments. **g**, IPCs partition into three subpopulations according to their response to glucose with or without sNPF receptor; see Methods. **h**, **i**, Individual traces of IPCs in control flies (**h**) and in flies in which *sNPF* receptor was rendered non-functional by expressing a dominant negative allele of *sNPF* receptor in the IPCs (**i**) responding to D-glucose. ***$P < 0.001$; unpaired two-tailed *t*-test (**c**). See Supplementary Table 1 for the sample sizes and statistical analyses.




**Extended Data Fig. 9 | *Crz* is not required in CN neurons for the two-choice behaviour and sugar-evoked Crz secretion. a**, **b**, Expression of GFP in the brains of flies carrying *Crz-Gal4, UAS-mCD8::GFP* (**a**) or *Crz receptor (CrzR)-Gal4, UAS-mCD8::GFP* (**b**). Scale bar, 100 µm. **c**, **d**, Knockdown of *Crz* or *Crz* receptor in flies carrying *Crz-Gal4* and *UAS-Crz* RNAi (**c**) or *CrzR-Gal4* and *UAS-CrzR* RNAis (**d**), respectively, does not impair the selection of D-glucose in starved flies. **e**, Immunoreactivity of intracellular Crz in CN neurons, probed with anti-Crz antibody, when the brains were incubated in 80 mM sucrose, 80 mM D-glucose (D-Glc), 80 mM D-glucose mixed with 0.5 µM TTX (D-Glc/TTX) or 80 mM L-glucose (L-Glc) in AHL. Scale bar, 5 µm. Images are *z*-stacked projections. One-way ANOVA with Tukey post hoc test (**c**, **e**) or unpaired two-tailed *t*-test (**d**). See Supplementary Table 1 for the sample sizes and statistical analyses.




**Extended Data Fig. 10 | Activity of IPCs was stimulated by sNPF application, and the circulating trehalose level in which CN neurons had been inactivated was reduced. a**, **b**, Average GCaMP traces (**a**) and Δ*F*/*F* (max) quantifications (**b**) from the IPCs in response to 80 μM sNPF in AHL or DMSO in AHL. **c**, An increase in circulating trehalose level in fed or starved flies in which CN neurons were inactivated. The levels of circulating trehalose in flies carrying *CN-Gal4* and *UAS-Kir2.1* and its heterozygote controls (*CN-Gal4*/+ and *UAS-Kir2.1*/+). **d**, A working model of the anatomic connectivities between CN neurons and IPCs (green), and between CN neurons and AKH-producing cells (red). CN neurons regulate glucose homeostasis by counter-balancing the activities of IPCs and AKH-producing cells through sNPF neurotransmitter that activates IPCs and inactivates AKH-producing cells. $^{*}P < 0.05$, $^{**}P < 0.01$ and $^{***}P < 0.001$; unpaired two-tailed *t*-test (**b**) and one-way ANOVA with Tukey post hoc test (**c**). See Supplementary Table 1 for the sample sizes and statistical analyses.

nature research

Corresponding author(s): Greg S.B. Suh

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

## Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☒ | ☐ | A description of all covariates tested |
| ☐ | ☒ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ | For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☒ | ☐ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒ | ☐ | Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |
| ☐ | ☒ | Clearly defined error bars<br>*State explicitly what error bars represent (e.g. SD, SE, CI)* |

*Our web collection on statistics for biologists may be useful.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | We did not use any unpublished code to collect the data in this study. We used LSM800 and ZEN software (Carl Zeiss, ZEN 2.3 SP1 FP1, version: 14.0.12.201) and Prairie two-photon microscope and its software (Prairie Technologies Inc. Prairie view v4.3.2.18). We also conducted mass spectrometry and dot blot assay as described in Methods. We also used Fiji software to collect dot blot images. |
| Data analysis | We did not use any unpublished code to analyze the data in this study. We used ZEM image analyzing software (Carl Zeiss, ZEN 2.3 SP1 FP1, version: 14.0.12.201), Graphpad prism 8.1.1, ImageJ 1.52a, Microsoft Excel (Microsoft office professional Plus 2016), and specific programs to analyze mass spectrometry data as described in method section (MaxQuant proteomics software v1.5.7.0, Thermo Scientific Xcalibur v4.1.31.9, Skyline Proteomics software v4.10.18169). |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All other raw data are available from the corresponding author on reasonable request. Raw mass spectrometry files have been deposited in the MassIVE database (https://massive.ucsd.edu/ProteoSAFe/static/massive.jsp); with MassIVE accession ID: MSV000083796.

# Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Sample size | No sample size was decided before the experiments. Sample size was determined based on the consistency of measurable differences between groups. We followed the previous study (Dus et al.,2015) to determine the statistical methods. We described the number of independent replications and sample size in "Statistics and reproducibility" section and Supplementary Table 1. |
| Data exclusions | In principle, we did not exclude any data in this study. However some trials were excluded because tested flies stopped moving during the behavioral assayes or tested cells were sick or died during the imaging experiments. |
| Replication | We replicated all the experiments and compared the data at least twice independently. |
| Randomization | We tried to randomize the population of flies in the same genotype. We matched the sex and age of flies in every trial. |
| Blinding | To reduce the bias, investigators checked genotypes after conducting the experiments and data collections. Key experiments were carried by multiple authors. |

# Reporting for specific materials, systems and methods

### Materials & experimental systems

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ Unique biological materials |
| ☐ | ☒ Antibodies |
| ☒ | ☐ Eukaryotic cell lines |
| ☒ | ☐ Palaeontology |
| ☐ | ☒ Animals and other organisms |
| ☒ | ☐ Human research participants |

### Methods

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ ChIP-seq |
| ☒ | ☐ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |

## Antibodies

Antibodies used

The primary antibodies used as follows: chicken anti-GFP (1:500; Invitrogen, A10262), rabbit anti-GFP (1:500; Invitrogen, A-11122), mouse anti-GFP (1:100; Sigma, G6539, use for synaptobrevin-GRASP), mouse anti-nc82 (1:25; Development Studies Hybridoma Bank, DSHB, AB-2314866), rabbit anti-dsRed (1:500; Clontech, 632496), rabbit anti-corazonin (Crz) (1:500; a gift from Jan Veenstra, Université de Bordeaux, France), rabbit anti-sNPF (1:500; a gift from Jan Veenstra, Université de Bordeaux, France), rabbit anti-dilp2 (1:500; a gift from Ernst Hafen, Institute for Molecular Systems Biology, Zürich, Switzerland), mouse anti-HA (1:500; Covance, 901501), and rabbit anti-AKH (1:500; gifts from Jae H. Park, University of Tennessee, Knoxville, TN, and Seung K. Kim, Stanford University, CA) antibodies. Secondary antibodies used as follows: Alexa Fluor 633 goat anti-rabbit IgG (1:500; Invitrogen, A-21070), Alexa Fluor 555 goat anti-mouse IgG (1:500; Invitrogen, A-21127), Alexa Fluor 555 goat anti-rabbit IgG (1:500; Invitrogen, A27039), Alexa Fluor 488 goat anti-rabbit IgG (1:500; Invitrogen, A27034), Alexa Fluor 488 goat anti-mouse

IgG (1:500; Invitrogen), and Alexa Fluor 488 goat anti-chicken IgG (1:500; Invitrogen, A28175).

Validation

All the primary antibodies used in this study were confirmed in the previous studies as listed below.
1: chicken anti-GFP (Invitrogen, A10262): Potdar, S. & Sheeba, V. Wakefulness Is promoted during day time by PDFR Signalling to dopaminergic neurons in Drosophila melanogaster. eneuro 5, ENEURO.0129-0118.2018 (2018).
2: rabbit anti-GFP (Invitrogen, A-11122): Zhang, P. et al. Heparan sulfate organizes neuronal synapses through neurexin partnerships. Cell 174, 1450-1464.e1423 (2018).
3: mouse anti-GFP (Sigma, G6539): Fushiki, A. et al. A circuit mechanism for the propagation of waves of muscle contraction in Drosophila. eLife 5, e13253 (2016).
4: mouse anti-nc82 (DSHB, AB-2314866): Ding, Y. et al. Neural evolution of context-dependent fly song. Curr. Biol. 29, 1089-1099.e1087 (2019).
5: rabbit anti-dsRed (Clontech, 632496): Ni, J. D. et al. Differential regulation of the Drosophila sleep homeostat by circadian and arousal inputs. eLife 8, e40487 (2019).
6: rabbit anti-corazonin (Crz) (Jan Veenstra): Kapan, N., Lushchak, O. V., Luo, J. & Nässel, D. R. Identified peptidergic neurons in the Drosophila brain regulate insulin-producing cells, stress responses and metabolism by coexpressed short neuropeptide F and corazonin. Cell. Mol. Life Sci. 69, 4051-4066 (2012).
7: rabbit anti-sNPF (Jan Veenstra): Knapek, S., Kahsai, L., Winther, Å. M. E., Tanimoto, H. & Nässel, D. R. Short neuropeptide F acts as a functional neuromodulator for olfactory memory in kenyon cells of Drosophila mushroom bodies. The Journal of Neuroscience 33, 5340-5345 (2013).
8: rabbit anti-dilp2 (Ernst Hafen): Ikeya, T., Galic, M., Belawat, P., Nairz, K. & Hafen, E. Nutrient-dependent expression of insulin-like peptides from neuroendocrine cells in the CNS contributes to growth regulation in Drosophila. Curr. Biol. 12, 1293-1300 (2002).
9: mouse anti-HA (Covance, 901501): Kim, Y. et al. Methylation-dependent regulation of HIF-1α stability restricts retinal and tumour angiogenesis. Nature Communications 7, 10347 (2016).
10: rabbit anti-AKH (Jae H. Park): Lee, G. & Park, J. H. Hemolymph sugar homeostasis and starvation-induced hyperactivity affected by genetic manipulations of the adipokinetic hormone-encoding gene in Drosophila melanogaster. Genetics 167, 311-323 (2004).
11: rabbit anti-AKH (Seung K. Kim): Kim, S. K. & Rulifson, E. J. Conserved mechanisms of glucose sensing and regulation by Drosophila corpora cardiaca cells. Nature 431, 316-320 (2004).

## Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals

Male flies (Drosophila melanogaster) between 3-10 days of age were used in this study. All the flies were outcrossed using w1118 (BL6326) over five times before the experiments

Wild animals

We did not use wild animals in this study. Instead, we used w1118 as a control.

Field-collected samples

We did not use any field-collected samples or animals in this study.

Article

# VISTA is an acidic pH-selective ligand for PSGL-1

https://doi.org/10.1038/s41586-019-1674-5

Received: 25 October 2018

Accepted: 20 September 2019

Published online: 23 October 2019

Robert J. Johnston[1]*, Linhui Julie Su[2,15], Jason Pinckney[3,15], David Critton[4,15], Eric Boyer[1], Arathi Krishnakumar[5], Martin Corbett[6], Andrew L. Rankin[7], Rose Dibella[2], Lynne Campbell[2], Gaelle H. Martin[8], Hadia Lemar[1], Thomas Cayton[1], Richard Y.-C. Huang[9], Xiaodi Deng[10], Akbar Nayeem[4], Haibin Chen[10], Burce Ergel[6], Joseph M. Rizzo[6], Aaron P. Yamniuk[6], Sanjib Dutta[3], Justine Ngo[11], Andrea Olga Shorts[11], Radha Ramakrishnan[9], Alexander Kozhich[9], Jim Holloway[12], Hua Fang[13], Ying-Kai Wang[13], Zheng Yang[9], Kader Thiam[8], Ginger Rakestraw[2], Arvind Rajpal[10], Paul Sheppard[1], Michael Quigley[1], Keith S. Bahjat[1] & Alan J. Korman[1,14]

Co-inhibitory immune receptors can contribute to T cell dysfunction in patients with cancer[1,2]. Blocking antibodies against cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) and programmed cell death 1 (PD-1) partially reverse this effect and are becoming standard of care in an increasing number of malignancies[3]. However, many of the other axes by which tumours become inhospitable to T cells are not fully understood. Here we report that V-domain immunoglobulin suppressor of T cell activation (VISTA) engages and suppresses T cells selectively at acidic pH such as that found in tumour microenvironments. Multiple histidine residues along the rim of the VISTA extracellular domain mediate binding to the adhesion and co-inhibitory receptor P-selectin glycoprotein ligand-1 (PSGL-1). Antibodies engineered to selectively bind and block this interaction in acidic environments were sufficient to reverse VISTA-mediated immune suppression in vivo. These findings identify a mechanism by which VISTA may engender resistance to anti-tumour immune responses, as well as an unexpectedly determinative role for pH in immune co-receptor engagement.

VISTA (also known as B7-H5, PD-1H, Gi24, Dies1, SISP1 and DD1α) is a B7 family ligand that is expressed on circulating and intratumoural myeloid cells and weakly expressed on activated lymphocytes[4–7]. It has been shown to inhibit T cell responses in vitro and in preclinical models of autoimmunity and cancer[4–14]. VISTA has also been recognized as a potential mediator of resistance to anti-PD-1 and anti-CTLA-4 immunotherapies in patients[6,7,15]. However, opportunities for therapeutic intervention have been limited by a lack of understanding of VISTA's counter-receptor and function. Here we report that VISTA is an acidic pH-selective ligand for PSGL-1.

Compared with other immunoglobulin superfamily members, the extracellular domain of VISTA is uniquely rich in histidine residues (Fig. 1a, Extended Data Fig. 1a). Because the imidazole sidechain of histidine protonates at physiologically relevant pH[16], we hypothesized that VISTA preferentially engages its counter-receptor in acidic environments such as tumour beds, where pH values as low as 5.85 have been measured[17,18]. We found that VISTA multimers bound detectably to leukocytes at acidic pH, but not at the physiological pH 7.4 (Fig. 1b, c, Extended Data Fig. 1b–d). Mouse VISTA also demonstrated pH-selective binding (Extended Data Fig. 1e). We tested a panel of monoclonal antibodies against VISTA for their ability to inhibit binding. VISTA.4 and other antibodies in the same epitope bin blocked VISTA binding to T cells at acidic pH, while VISTA.5 and other antibodies in its epitope bin did not (Fig. 1d, Extended Data Fig. 1f–h). VISTA has been shown to inhibit T cell function in a variety of contexts and without a clear requirement for acidic pH[4,6,9,11,13]. To address the role of pH in VISTA function, we cultured T cells and Jurkat cells with VISTA-expressing cells or recombinant VISTA. Whereas VISTA suppressed proliferation, IFN-γ production, and NF-κB phosphorylation at pH 7.4, its effect was much more pronounced at acidic pH (Fig. 1e, f, Extended Data Fig. 2a–e). VISTA.4, but not VISTA.5, restored T cell responsiveness (Fig. 1e, f, Extended Data Fig. 2a–e). These results suggested that VISTA is functionally pH-selective, and that antibodies that block VISTA binding to T cells at acidic pH reverse its suppressive activity.

We subsequently found that VISTA.4 recognizes a histidine-rich epitope and is itself sensitive to pH in binding assays (Extended Data Fig. 2f–i). This led us to hypothesize that histidine protonation enables antibodies to distinguish the active (acidic pH) and inactive (neutral pH) states of the VISTA ligand interface. Mutational scanning of VISTA.4 identified variants with a broad range of pH preferences (Fig. 1g, Extended Data Fig. 3a). Further engineering of rare non-pH-sensitive blocking antibodies such as VISTA.16 produced progeny that were

[1]Immuno-Oncology Discovery, Bristol-Myers Squibb, Redwood City, CA, USA. [2]Selection Technologies and Protein Engineering, Bristol-Myers Squibb, Cambridge, MA, USA. [3]Biomolecular Characterization, Bristol-Myers Squibb, Cambridge, MA, USA. [4]Molecular Structure and Design, Bristol-Myers Squibb, Lawrenceville, NJ, USA. [5]Proteomics Discovery, Bristol-Myers Squibb, Hopewell, NJ, USA. [6]Protein Sciences, Bristol-Myers Squibb, Lawrenceville, NJ, USA. [7]Immuno-Oncology Discovery, Five Prime Therapeutics, South San Francisco, CA, USA. [8]genOway, Lyon, France. [9]Pharmaceutical Candidate Optimization, Bristol-Myers Squibb, Lawrenceville, NJ, USA. [10]Protein Therapeutics and Biologics Lead Discovery, Bristol-Myers Squibb, Redwood City, CA, USA. [11]Translational Sciences, Bristol-Myers Squibb, Lawrenceville, NJ, USA. [12]Translational Bioinformatics, Bristol-Myers Squibb, Redwood City, CA, USA. [13]Lead Discovery and Optimization, Bristol-Myers Squibb, Lawrenceville, NJ, USA. [14]Present address: Vir Biotechnology, San Francisco, California, USA. [15]These authors contributed equally: Linhui Julie Su, Jason Pinckney, David Critton. *e-mail: robert.johnston@bms.com

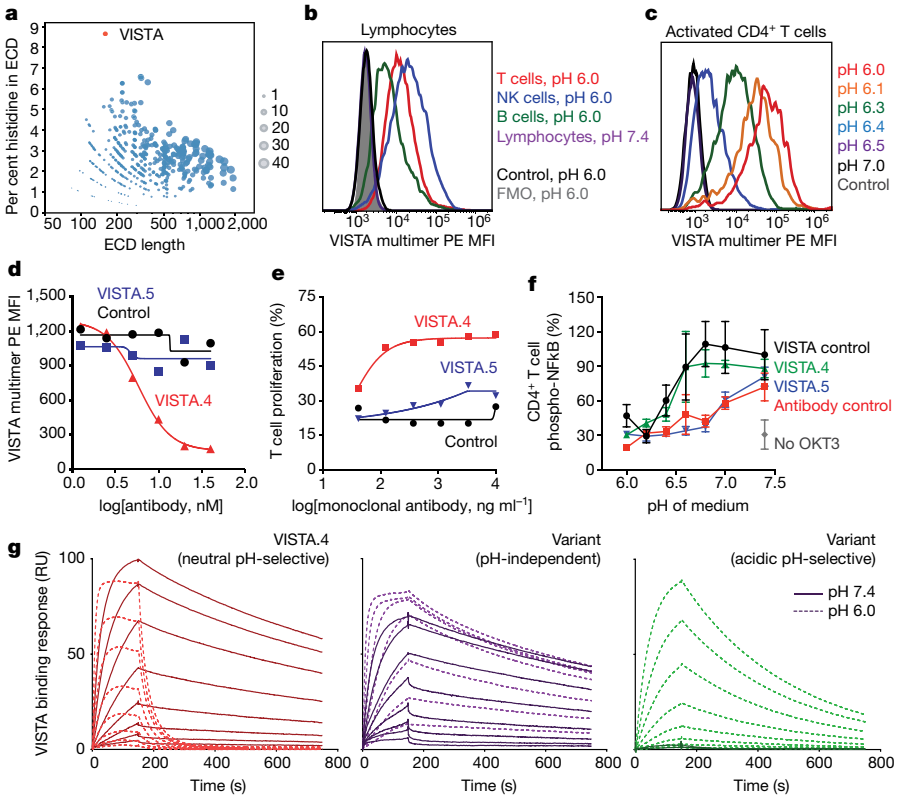



**Fig. 1 | VISTA is pH-selective. a**, Immunoglobulin superfamily extracellular domain histidine frequencies and lengths. The size of each point corresponds to the number of histidine residues. VISTA is highlighted in red. **b**, VISTA multimer binding to human lymphocytes at pH 6.0 and 7.4. Lymphocytes left unbound (FMO) or bound with non-VISTA-loaded multimers are included as controls. These data are representative of three independent experiments. **c**, VISTA multimer binding to activated human CD4$^+$ T cells at the indicated pH. These data are representative of seven independent experiments. **d**, VISTA multimer binding to activated human CD4$^+$ T cells at pH 6.0 and in the presence of VISTA.4 (red), VISTA.5 (blue) or a control antibody (black). Data are VISTA multimer mean fluorescence intensity (MFI) and are representative of six independent experiments. **e**, Effects of VISTA.4 (red), VISTA.5 (blue) and control antibodies (black) on human CD4$^+$ T cell proliferation during co-culture with 293T-OKT3-VISTA cells, which ectopically express a single chain variable fragment of the agonistic CD3 monoclonal antibody OKT3 and human VISTA. Data are the per cent of T cells that proliferated and are representative of three independent experiments. **f**, Effects of VISTA.4 (green), VISTA.5 (blue) and control antibodies (red) on human CD4$^+$ T cell NF-kB phosphorylation during T cell receptor and VISTA stimulation at the indicated pH. Non-T cell receptor-stimulated cells (grey) and non-VISTA-stimulated cells (black) are also included as controls. Data are pNF-κB MFI normalized to control ± s.e.m. $n = 2$ T cell donors; these data are representative of two independent experiments. **g**, Surface plasmon resonance (SPR) sensorgrams of VISTA.4, a pH-independent variant of VISTA.4 and an acidic pH-selective variant of VISTA.4 binding to human VISTA at pH 6.0 (dashed traces) and pH 7.4 (solid traces). These data are representative of two independent experiments.

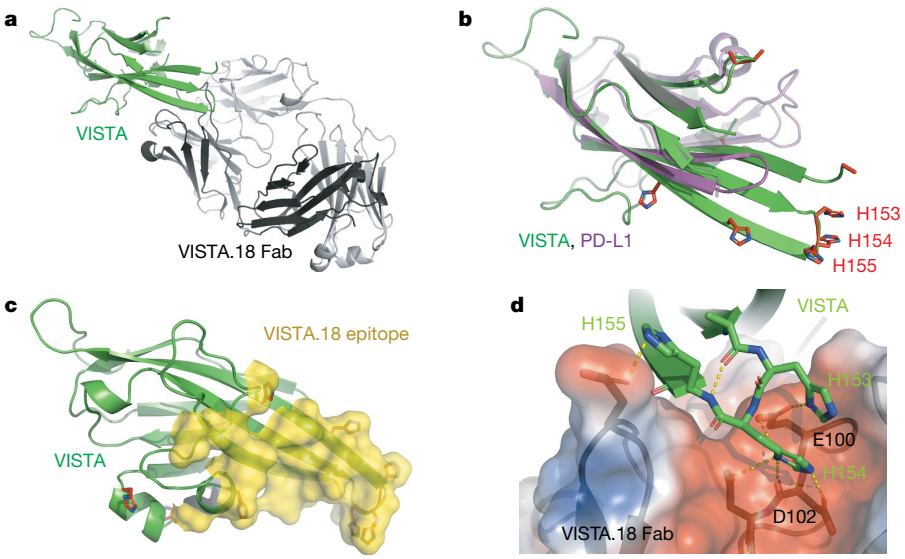



**Fig. 2 | Crystal structure of VISTA and blocking antibody epitope. a**, The structure of the human VISTA IgV domain (green) in complex with the VISTA.18 Fab (heavy chain, dark grey; light chain, light grey). **b**, A superimposition of the VISTA (green) and PD-L1 (purple) IgV domains. VISTA histidine residues are depicted in stick representation. Histidine residues occupying the loop between the central β-sheets of VISTA (H153, H154 and H155) are labelled. H100, H101 and H104 are in disordered regions and are not depicted. **c**, The molecular surface of the VISTA.18 epitope (yellow). **d**, An enlarged view of the interface between VISTA (green, with epitope residues depicted in stick representation) and VISTA.18 (depicted as an electrostatic surface).

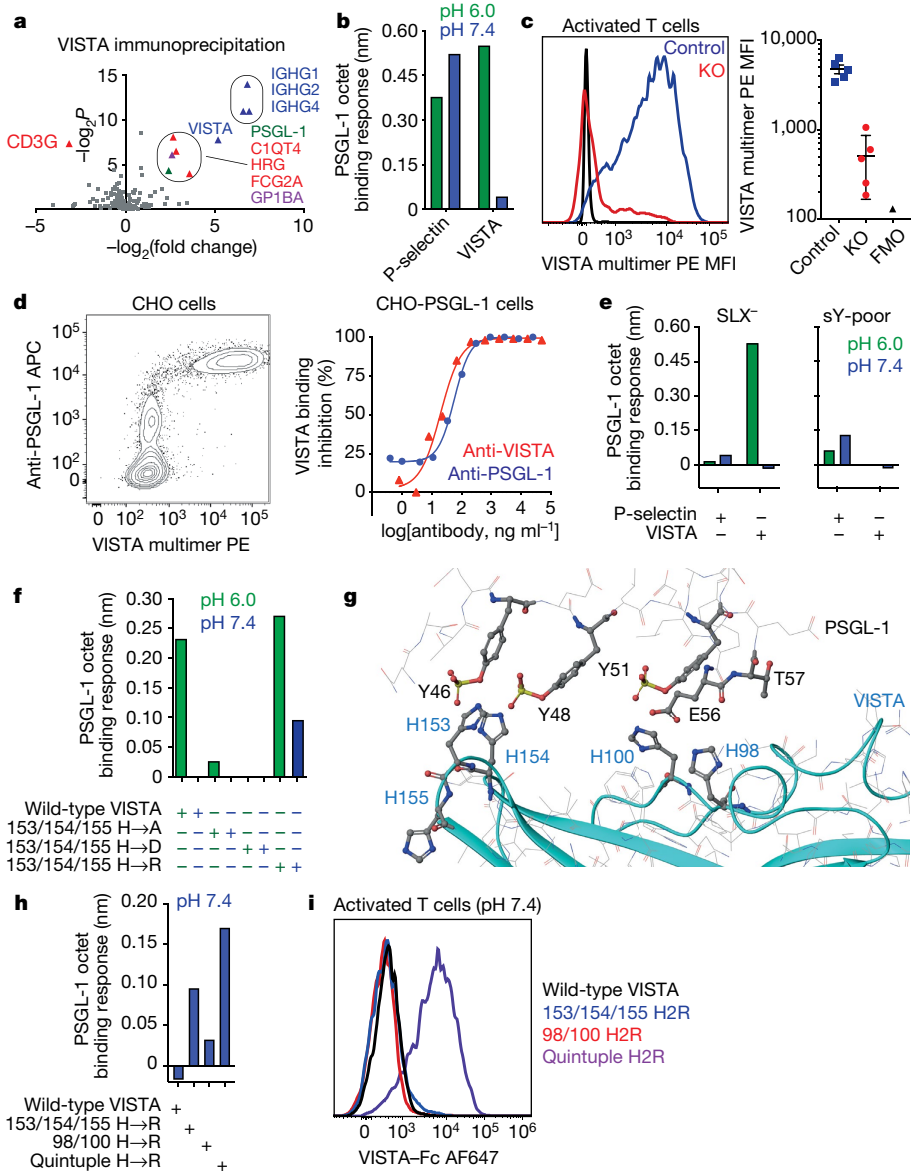



**Fig. 3 | PSGL-1 is a VISTA receptor at acidic pH. a**, Results of VISTA–Fc (right) and anti-CD3 (left) receptor capture on human CD4$^+$ T cells at pH 6.0. $\log_2$(fold enrichment) and −log adjusted *P* values are plotted on the *x* and *y* axes, respectively. Among the captured proteins, bait components are coloured blue and putative counter-receptors are coloured red, green (VISTA) and purple (GP1BA). These data are representative of two independent experiments. **b**, Bio-layer interferometry (BLI) binding magnitudes for P-selectin and VISTA binding to captured PSGL-1 at pH 6.0 (green) and 7.4 (blue). These data are representative of ten independent experiments. **c**, VISTA multimer binding to activated human CD4$^+$ T cells at pH 6.0 with (red) and without (blue) PSGL-1 gene deletion. Unbound cells (FMO, black) are included as controls. Data are VISTA multimer MFI ± s.e.m. and are a composite of five independent experiments. **d**, Left, VISTA multimer binding to CHO cells expressing PSGL-1 at pH 6.0. Right, effects of VISTA.4 (red) and the PSGL-1 antibody KPL1 (blue) on binding. Data are the per cent reduction of VISTA-Fc MFI relative to control and are representative of eight independent experiments. **e**, BLI binding magnitudes of VISTA and P-selectin binding to PSGL-1 produced without sialyl-Lewis X decoration (SLX$^-$) or with low levels of tyrosine sulfation (sY-poor) at pH 6.0 (green) and pH 7.4 (blue). These data are representative of one experiment. **f**, BLI binding magnitudes of PSGL-1 binding to VISTA with histidine residues at positions 153/154/155 left intact (WT) or replaced by alanine (H→A), aspartic acid (H→D), or arginine (H→R) at pH 6.0 (green) and pH 7.4 (blue). These data are representative of three independent experiments. **g**, A computational model of PSGL-1 (grey) bound to VISTA (cyan). Key residues are depicted in stick representation. **h**, BLI magnitudes for PSGL-1 binding to wild-type, H98R/H100R, H153R/H154R/H155R and H98R/H100R/H153R/H154R/H155R (quintuple) VISTA at pH 7.4. These data are representative of two independent experiments. **i**, Wild-type and H→R mutant VISTA–Fc binding to activated human T cells at pH 7.4. These data are representative of two independent experiments. WT, wild-type.

highly selective for acidic pH, including the clone VISTA.18 (Extended Data Fig. 3b–d).

To characterize VISTA's structure and the determinants of blocking-antibody binding, we co-crystallized the VISTA immunoglobulin variable (IgV) domain with the fragment antigen-binding (Fab) domain of VISTA.18 (Fig. 2a, Extended Data Fig. 4a, Supplementary Table 1). The VISTA IgV domain is elongated by additional residues in its two C-terminal β-strands, and histidine residues are concentrated on a loop connecting the strands (Fig. 2b). VISTA.18 binds this loop, whereas the non-blocking antibody VISTA.5 binds an opposing region (Fig. 2c, Extended Data Fig. 4b, c). The VISTA.18 complementarity-determining region (CDR) residues E100 and D102 form hydrogen bonds with VISTA histidine residues H153 and H154 and mediate the antibody's selectivity for acidic pH (Fig. 2d Extended Data Fig. 4d, e).

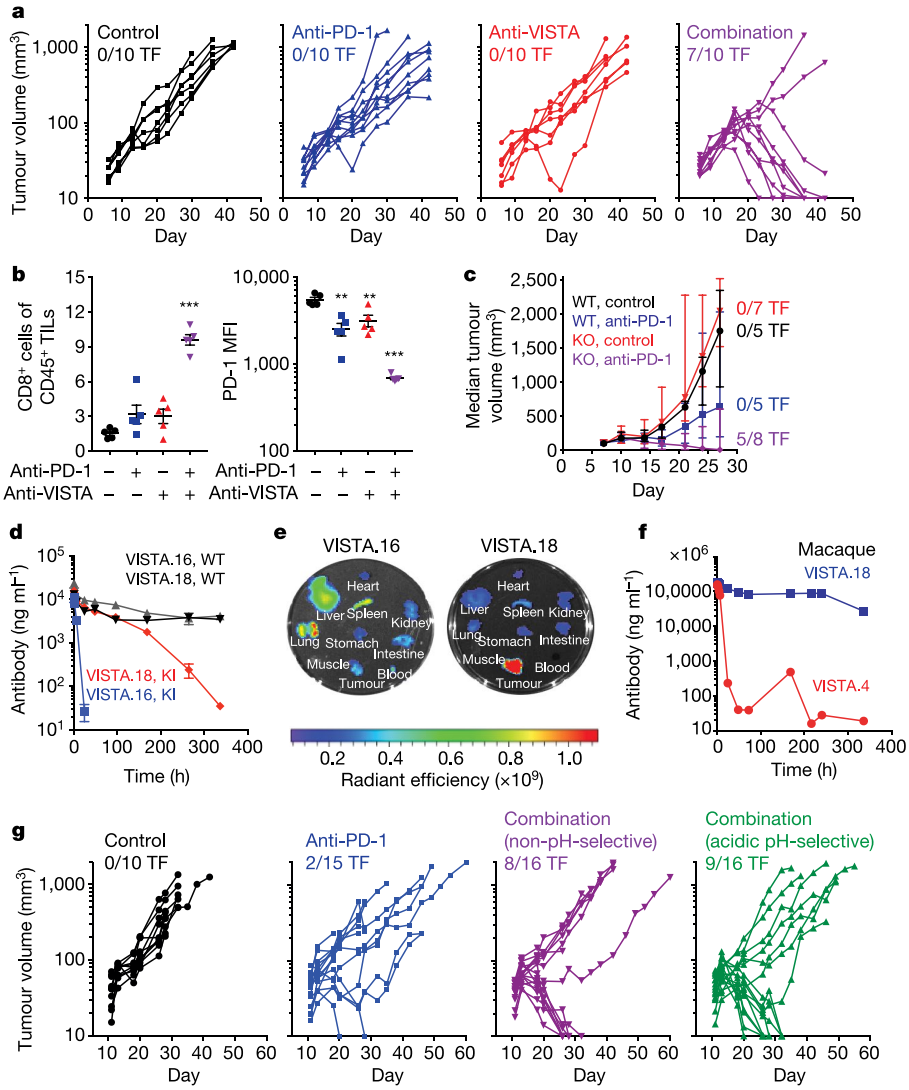



**Fig. 4 | VISTA blockade at acidic pH reverses its suppressive effects in vivo.** **a**, **b**, MC38 tumour-bearing wild-type mice were treated with a control antibody, a VISTA-blocking antibody, a PD-1-blocking antibody, or both VISTA- and PD-1-blocking antibodies. $n = 10$ (**a**) or 5 (**b**) mice per group; these data are representative of three independent experiments. **a**, Tumour volumes over time. TF denotes mice that were tumour-free at the end of the study. **b**, Frequency (left) and PD-1 MFI (right) of intratumoural $CD8^+$ T cells. Data are means ± s.e.m. with one-way ANOVA and Dunnett's multiple comparisons. $^{\#}P = 0.001$, $^{\#\#}P = 0.007$ and $^{***}P < 0.0001$. **c**, MC38 tumour-bearing VISTA-knockout (KO) mice and wild-type littermates were treated with a control antibody (VISTA knockout, red; wild type, black) or a PD-1 blocking antibody (VISTA knockout, purple; wild type, blue). $n = 5$–8 as indicated; these data are representative of two independent experiments. Data are medians ± interquartile ranges. **d**, Human VISTA knock-in (KI) mice and their wild-type littermates were treated with VISTA.16 (VISTA knock-in, blue; wild type, black) or VISTA.18 (VISTA knock-in, red; wild type, grey). Data are mean blood antibody concentrations ± s.e.m. The calculated MRT for VISTA.16 and VISTA.18 in VISTA knock-in mice were 4.1 and 71 h, respectively. $n = 4$ knock-in; $n = 1$ VISTA.16, wild-type; and $n = 2$ VISTA.18, wild-type mice per group; data are representative of one experiment. **e**, Human VISTA knock-in mice bearing MC38 tumours were treated with fluorescently labelled VISTA.16 (left) or VISTA.18 (right). Organs were imaged at 51 h post-injection. These data are representative of two independent experiments. **f**, Cynomolgus macaques were treated with VISTA.4 (red circles) or VISTA.18 (blue squares). Data are serum antibody concentrations. The calculated MRTs were 7.6 h and 717 h respectively. $n = 1$ macaque per antibody; these data are representative of one experiment. **g**, Human VISTA knock-in mice bearing MC38 tumours were treated with a control antibody, a PD-1 blocking antibody, the non-pH-selective antibody VISTA.16 and/or the acidic pH-selective antibody VISTA.18. Data are tumour volumes. $n = 10$–16 per group as indicated; these data are a composite of two independent experiments.

We then performed ligand-based receptor capture with VISTA–Fc chimeric protein and human $CD4^+$ T cells at acidic pH[19]. PSGL-1 was one of few proteins that were enriched relative to controls (Fig. 3a). In addition to its well-characterized role facilitating adhesion interactions between leukocytes, platelets and endothelial cells[20,21], PSGL-1 has been identified as a negative regulator of T cell responses in contexts of chronic viral infection, cancer, and some autoimmune diseases[22–28]. To confirm the interaction, we conducted Octet biosensor and isothermal titration calorimetry assays with the minimal PSGL-1 glycopeptide that supports high affinity P-selectin binding[29] (Extended Data Fig. 5). PSGL-1 bound VISTA selectively at acidic pH with an affinity of 0.65–0.85 μM and a stoichiometry approaching 1:1 (Fig. 3b, Extended Data Fig. 6a–d, Supplementary Table 2). Binding was blocked by recombinant P-selectin, VISTA blocking antibodies and the P-selectin-blocking PSGL-1 antibody KPL1 (Extended Data Fig. 6e–h). Gene deletion of PSGL-1 significantly reduced VISTA binding to T cells (Fig. 3c, Extended Data Fig. 6i), whereas ectopic expression of PSGL-1 was sufficient to enable VISTA binding to Chinese hamster ovary cells at acidic pH (Fig. 3d and Extended Data Fig. 6j). Similar to P-selectin[30], VISTA also bound modestly to heparan sulfate (Extended Data Fig. 6k).

Glycoprotein 1b platelet subunit-α (GP1BA) was captured by VISTA as well (Fig. 3a), but we could not detect VISTA binding to recombinant GP1BA or to platelets (Extended Data Fig. 7a, b). VISTA has recently been reported to bind V-set immunoglobulin domain containing 3 (VSIG-3), a surface receptor expressed in brain, testis and some cancer tissues[31,32]. We observed moderately pH-selective binding of VISTA to VSIG-3 using the Octet biosensor, but could not detect specific binding in cell-based assays, and found no competition between VSIG-3 and PSGL-1 (Extended Data Fig. 7c–h). VISTA has also been reported to engage in homotypic binding[33], but we were unable to confirm this interaction (Extended Data Fig. 7i).

We next characterized the specificity of PSGL-1–VISTA binding. PSGL-1 binding to P-selectin is supported by sulfotyrosine and sialyl-Lewis X tetrasaccharide post-translational modifications[29]. VISTA binding to PSGL-1 was independent of sialyl-Lewis X but dependent on tyrosine sulfation (Fig. 3e, Extended Data Figs. 5d, 8a, b).

Blocking antibody coverage of VISTA histidine residues H153, H154 and H155 suggested that these residues support PSGL-1 binding (Fig. 2). Replacement of these histidines with negatively charged aspartic acid eliminated VISTA binding and function in biophysical and cell-based assays (Fig. 3f, Extended Data Fig. 8c–f). Replacement with positively charged arginine left VISTA activity at acidic pH intact, but conferred only weak binding at pH 7.4, indicating that these three histidines were necessary but not sufficient for binding (Fig. 3f, Extended Data Fig. 8c–f). To identify other relevant residues, we used the solved structures of PSGL-1 bound to P-selectin[34] and VISTA bound to VISTA.18 Fab (Fig. 2) to computationally model PSGL-1 docked to VISTA (Fig. 3g). In this model, sulfated PSGL-1 tyrosine residues Y46 and Y48 make ionic interactions with protonated VISTA histidine residues H153 and H154, whereas VISTA H98 and H100 appear to interact with PSGL-1 E56 and Y51. Substitution of histidines H98, H100, H153, H154 and H155 with arginine enabled VISTA to bind to PSGL-1 and to T cells at pH 7.4 (Fig. 3h, i, Extended Data Fig. 8g). In the same model, the hydroxyl group of PSGL-1 T57, which can be decorated with sialyl-Lewis X, points away from VISTA, consistent with the negligible influence of sialyl-Lewis X on VISTA–PSGL-1 binding (Fig. 3g). These data demonstrated that PSGL-1 binding to VISTA is mediated by charged interactions between sulfated tyrosine and protonated histidine residues.

Finally, we examined the role of VISTA in anti-tumour immune responses. Whereas single agent treatment with a mouse VISTA blocking antibody had little effect, co-blockade of VISTA and PD-1 elicited tumour rejection in a majority of mice implanted with MC38 colorectal adenocarcinomas (Fig. 4a, Extended Data Fig. 9a). Combination therapy also enhanced T cell tumour infiltration and reduced intratumoural T cell expression of the co-inhibitory receptors PD-1, LAG-3 and TIM-3 (Fig. 4b, Extended Data Fig. 9b, c). Intratumoural myeloid cell frequencies were largely unaffected by VISTA blockade (Extended Data Fig. 9d). These results were phenocopied in VISTA-knockout mice, indicating that blocking antibodies can reverse VISTA-mediated immune suppression (Fig. 4c, Extended Data Fig. 9e).

To assess VISTA-mediated suppression within tumour microenvironments, we treated mice expressing the human VISTA extracellular domain (Extended Data Fig. 9f–h) with acidic pH-selective and non-pH-selective human VISTA-blocking antibodies. VISTA expression on circulating and organ-resident myeloid cells subjects antibodies to extensive target-mediated drug disposition (Extended Data Fig. 9i). Consistent with this effect, the non-pH-selective antibody VISTA.16 exhibited a short blood mean residence time (MRT) and localized primarily to leukocyte-rich organs (Fig. 4d, e, Extended Data Fig. 10a). By contrast, the acidic pH-selective antibody VISTA.18 accumulated preferentially within tumours and exhibited a much longer blood and serum MRT in mice and a cynomolgus macaque, confirming its inability to engage VISTA efficiently at physiological pH in vivo (Fig. 4d–f, Extended Data Fig. 10a–e). VISTA.18 nevertheless matched VISTA.16 in therapeutic benefit in combination with anti-PD-1 (Fig. 4g, Extended Data Fig. 10f). These data support the hypothesis that VISTA functions as an acidic-pH-selective immune checkpoint, although further study is needed into its mechanisms of action and the relevance of PSGL-1 in vivo.

Typically, activation-induced expression of co-inhibitory receptors results in preferential restraint of maturing immune responses[35]. VISTA instead appears to utilize pH selectivity to achieve a similar outcome, with suppression occurring in inflamed and acidic environments such as tumours rather than in lymphoid organs or the blood. This suggests that immune responses can be regulated by checkpoints specific to acidic environments, and that further study of pH selectivity may afford new opportunities for immunotherapeutic drug development.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1674-5.


1. Thommen, D. S. & Schumacher, T. N. T cell dysfunction in cancer. *Cancer Cell* **33**, 547–562 (2018).
2. McLane, L. M., Abdel-Hakeem, M. S. & Wherry, E. J. CD8 T cell exhaustion during chronic viral infection and cancer. *Annu. Rev. Immunol.* **37**, 457–495 (2019).
3. Ribas, A. & Wolchok, J. D. Cancer immunotherapy using checkpoint blockade. *Science* **359**, 1350–1355 (2018).
4. Wang, L. et al. VISTA, a novel mouse Ig superfamily ligand that negatively regulates T cell responses. *J. Exp. Med.* **208**, 577–592 (2011).
5. Flies, D. B., Wang, S., Xu, H. & Chen, L. Cutting edge: A monoclonal antibody specific for the programmed death-1 homolog prevents graft-versus-host disease in mouse models. *J. Immunol.* **187**, 1537–1541 (2011).
6. Gao, J. et al. VISTA is an inhibitory immune checkpoint that is increased after ipilimumab therapy in patients with prostate cancer. *Nat. Med.* **23**, 551–555 (2017).
7. Blando, J. et al. Comparison of immune infiltrates in melanoma and pancreatic cancer highlights VISTA as a potential target in pancreatic cancer. *Proc. Natl Acad. Sci. USA* **116**, 1692–1697 (2019).
8. Le Mercier, I. et al. VISTA regulates the development of protective antitumor immunity. *Cancer Res.* **74**, 1933–1944 (2014).
9. Lines, J. L. et al. VISTA is an immune checkpoint molecule for human T cells. *Cancer Res.* **74**, 1924–1932 (2014).
10. Wang, L. et al. Disruption of the immune-checkpoint VISTA gene imparts a proinflammatory phenotype with predisposition to the development of autoimmunity. *Proc. Natl Acad. Sci. USA* **111**, 14846–14851 (2014).
11. Liu, J. et al. Immune-checkpoint proteins VISTA and PD-1 nonredundantly regulate murine T-cell responses. *Proc. Natl Acad. Sci. USA* **112**, 6682–6687 (2015).
12. Ceeraz, S. et al. VISTA deficiency accelerates the development of fatal murine lupus nephritis. *Arthritis Rheumatol.* **69**, 814–825 (2017).
13. Li, N. et al. Immune-checkpoint protein VISTA critically regulates the IL-23/IL-17 inflammatory axis. *Sci. Rep.* **7**, 1485 (2017).
14. Sergent, P. A. et al. Blocking the VISTA pathway enhances disease progression in (NZB×NZW) F1 female mice. *Lupus* **27**, 210–216 (2018).
15. Kakavand, H. et al. Negative immune checkpoint regulation by VISTA: a mechanism of acquired resistance to anti-PD-1 therapy in metastatic melanoma patients. *Mod. Pathol.* **30**, 1666–1676 (2017).
16. Schönichen, A., Webb, B. A., Jacobson, M. P. & Barber, D. L. Considering protonation as a posttranslational modification regulating protein structure and function. *Annu. Rev. Biophys.* **42**, 289–314 (2013).
17. Wike-Hooley, J. L., Haveman, J. & Reinhold, H. S. The relevance of tumour pH to the treatment of malignant disease. *Radiother. Oncol.* **2**, 343–366 (1984).
18. Kato, Y. et al. Acidic extracellular microenvironment and cancer. *Cancer Cell Int.* **13**, 89 (2013).
19. Frei, A. P. et al. Direct identification of ligand-receptor interactions on living cells and tissues. *Nat. Biotechnol.* **30**, 997–1001 (2012).
20. Carlow, D. A. et al. PSGL-1 function in immunity and steady state homeostasis. *Immunol. Rev.* **230**, 75–96 (2009).
21. Abadier, M. & Ley, K. P-selectin glycoprotein ligand-1 in T cells. *Curr. Opin. Hematol.* **24**, 265–273 (2017).
22. Matsumoto, M., Miyasaka, M. & Hirata, T. P-selectin glycoprotein ligand-1 negatively regulates T-cell immune responses. *J. Immunol.* **183**, 7204–7211 (2009).
23. Pérez-Frías, A. et al. Development of an autoimmune syndrome affecting the skin and internal organs in P-selectin glycoprotein ligand 1 leukocyte receptor-deficient mice. *Arthritis Rheumatol.* **66**, 3178–3189 (2014).
24. Tinoco, R. et al. PSGL-1 is an immune checkpoint regulator that promotes T Cell exhaustion. *Immunity* **44**, 1190–1203 (2016).
25. Urzainqui, A. et al. Functional role of P-selectin glycoprotein ligand 1/P-selectin interaction in the generation of tolerogenic dendritic cells. *J. Immunol.* **179**, 7457–7465 (2007).

26. Nuñez-Andrade, N. et al. P-selectin glycoprotein ligand-1 modulates immune inflammatory responses in the enteric lamina propria. *J. Pathol.* **224**, 212–221 (2011).
27. Veerman, K. M. et al. PSGL-1 regulates the migration and proliferation of $CD8^+$ T cells under homeostatic conditions. *J. Immunol.* **188**, 1638–1646 (2012).
28. Angiari, S. et al. Regulatory T cells suppress the late phase of the immune response in lymph nodes through P-selectin glycoprotein ligand-1. *J. Immunol.* **191**, 5489–5500 (2013).
29. Sako, D. et al. A sulfated peptide segment at the amino terminus of PSGL-1 is critical for P-selectin binding. *Cell* **83**, 323–331 (1995).
30. Nelson, R. M. et al. Heparin oligosaccharides bind L- and P-selectin and inhibit acute inflammation. *Blood* **82**, 3253–3258 (1993).
31. Wang, J. et al. VSIG-3 as a ligand of VISTA inhibits human T-cell function. *Immunology* **156**, 74–85 (2019).
32. Mehta, N. et al. Structure and functional binding epitope of V-domain Ig suppressor of T cell activation. *Cell Rep.* **28**, 2509–2516 (2019).
33. Yoon, K. W. et al. Control of signaling-mediated clearance of apoptotic cells by the tumor suppressor p53. *Science* **349**, 1261669 (2015).
34. Somers, W. S., Tang, J., Shaw, G. D. & Camphausen, R. T. Insights into the molecular basis of leukocyte tethering and rolling revealed by structures of P- and E-selectin bound to $SLe^X$ and PSGL-1. *Cell* **103**, 467–479 (2000).
35. Ahn, E. et al. Role of PD-1 during effector CD8 T cell differentiation. *Proc. Natl Acad. Sci. USA* **115**, 4749–4754 (2018).

# Methods

No statistical methods were used to predetermine sample size. Randomization on the basis of tumour volume was performed prior to the initiation of treatment in mouse tumor studies. The investigators were not blinded to allocation during experiments and outcome assessment.

### Immunoglobulin superfamily histidine analysis and VISTA sequence alignments

The amino acid sequences of the extracellular domains of immunoglobulin domain-containing proteins were extracted from the UniProt and Swiss-Prot databases. The number and frequency of histidine residues were calculated for each protein. NCBI VISTA amino acid reference sequences for *Homo sapiens* (NP_071436.1), *Pan troglodytes* (XP_001135701.2), *Macaca fascicularis* (predicted, XP_015311697.1), *Canis lupus familiaris* (XP_013968352.2), *Rattus norvegicus* (NP_001037765.1) and *Mus musculus* (NP_083008.1) were selected for sequence alignment. VISTA residues are numbered inclusive of the signal peptide.

### Cell binding and blocking assays

In human VISTA binding and blocking assays, phycoerythrin (PE)-conjugated streptavidin multimers (Klickmers, Immudex) were diluted to 32 nM in Hank's Buffered Salt Solution (HBSS, with calcium and magnesium) adjusted with 2-(*N*-morpholino)ethanesulfonic acid (MES) to the indicated pH. The diluted dextramers were loaded with 32–900 nM recombinant monobiotinylated human VISTA (ACRO Biosystems) to facilitate VISTA-dextramer capture. 'Empty' multimers that were not incubated with VISTA were used as a negative control. Human leukocytes, unstimulated peripheral blood mononuclear cells (PBMC) T cells, PBMC T cells stimulated for 72–96 h with anti-CD3/CD28 bead stimulation (Human T-activator Dynabeads, ThermoFisher), or Chinese Hamster Ovary (CHO) cells were labelled with VISTA multimers for 30 min at room temperature before washing with the same HBSS + MES buffers. Alternatively, cells were labelled with human VISTA–Fc chimeric proteins, and binding was detected with anti-IgG secondary antibodies (Jackson ImmunoResearch) diluted into the same HBSS + MES buffers. Labelled cells were left unfixed or fixed with formaldehyde (FoxP3 fixation buffer, eBioscience) and acquired on a flow cytometer.

In VISTA antibody blocking assays, 100 nM-loaded VISTA multimers or VISTA–Fc chimeric proteins were pre-incubated with the indicated antibodies before cell binding. In recombinant protein blocking assays, cells were pre-incubated with the indicated recombinant proteins before labelling with 100 nM-loaded VISTA multimers or VISTA–Fc chimeric proteins.

In PSGL-1 antibody blocking assays, cells were pre-incubated with KPL1 (BD Biosciences or Biolegend) or PL2 (MBL) before labelling with 32 nM-loaded VISTA multimers or VISTA–Fc chimeric proteins. VISTA–Fc binding was detected by anti-IgG (Jackson ImmunoResearch) or 6×His (Columbia Biosciences) antibodies. Cells were acquired by flow cytometry or homogenous time resolved fluorescence (HTRF).

In mouse VISTA binding and blocking assays, mouse splenocytes and lymph node-resident cells were used directly ex vivo or first stimulated for 48 h with anti-CD3/CD28 bead stimulation (Mouse T-activator Dynabeads, ThermoFisher). Cells were labelled with mouse VISTA–Fc chimeric proteins in pH 6.0 HBSS or PBS. VISTA–Fc, and binding was detected with anti-IgG secondary antibodies (Jackson ImmunoResearch). In antibody blocking assays, mouse VISTA–Fc was pre-incubated with VISTA.10 before cell labelling. Binding of VISTA to individual leukocyte subsets was determined by staining for CD4, CD8, B220 and CD11b (ThermoFisher). Cells were acquired on a flow cytometer.

In VSIG-3 binding assays, CHO and HEK293 cells were engineered to ectopically express human VSIG-3 and VISTA respectively. VSIG-3 expression was confirmed by flow cytometry using anti-VSIG-3 (pAb AF4915, R&D Systems). VISTA expression was confirmed by flow cytometry using anti-VISTA (clone 740804, R&D Systems). Cell binding assays were performed in PBS buffers containing 0.9 mM $CaCl_2$, 0.05 mM $MgCl_2$ and 0.5% BSA that were adjusted to the indicated pH by varying the ratios of $Na_2HPO_4$ and $KH_2PO_4$. VISTA–Fc and VSIG-3–Fc were used at 10 μg $ml^{-1}$. Binding was detected with anti-human IgG Fab'2–PE (Invitrogen).

### Recombinant VISTA and PSGL-1 proteins

Histidine-tagged human, cynomolgus macaque, and mouse VISTA extracellular domains were produced by transient transfection of Expi293 cells. Proteins were affinity purified via the histidine tag and then size-exclusion chromatography (SEC) (Superdex200). Human and mouse VISTA–Fc chimeric proteins were purchased from R&D Systems or produced by transient transfection of Expi293 cells. VISTA–Fc proteins with H98, H100, H153, H154, and H155 residues mutated to alanine, aspartic acid, and/or arginine were produced by transient transfection of Expi293 cells.

Human PSGL-1–Fc and P-selectin–Fc chimeric proteins were purchased from R&D Systems. PSGL-119-mer glycopeptides were produced as previously described[29]. In brief, the human PSGL-119-mer contained the 19 N-terminal residues of PSGL1 fused to a human IgG1 Fc via a (G)4S linker and a TVMV protease site (19-mer–Fc). To enable PSGL-1 decoration with sialyl-Lewis X (a tetrasaccharide containing sialic acid), PSGL-1 19-mer–Fc plasmid was co-expressed with and without the addition of plasmids encoding glucosaminyl (N-acetyl) transferase (core 2, GCNT1) and alpha (1,3)-fucosyltransferase-7 enzymes at an 8:1:1 ratio in Expi293 cells. The 19-mer–Fc fusion proteins were purified from supernatant by MabSelectSure Protein A resin (GE Healthcare) followed by preparative SEC (Superdex200, GE Healthcare). Where indicated, the 19-mer fusion proteins were further fractionated into sulfotyrosine-enriched and sulfotyrosine-depleted pools by separated on a Q HP (GE) column in Tris buffer at pH 7.5. The presence of the sialyl-Lewis X and tyrosine sulfation post-translational modifications were determined by HECA452 antibody binding enzyme-linked immunosorbency assay (ELISA), human P-selectin binding ELISA, and mass spectrometry peptide mapping.

Human VSIG-3–Fc chimeric proteins were purchased from R&D Systems.

### Antibody generation

Anti-human VISTA antibodies were produced in transgenic mice expressing human immunoglobulin alleles in place of murine alleles[36]. Mice were immunized with recombinant human VISTA. Splenocytes from these mice were fused with the Sp2/0 myeloma cell line, and fusions that were positive for human IgG g/κ antibody production were screened for VISTA reactivity by ELISA to his-tagged human VISTA and by flow cytometry to HEK293 cells stably expressing cell surface human VISTA. Alternatively, single chain fragment-variable (scFv) antibody libraries were created from genetic material isolated from immunized mice and screened by mRNA display for binding to recombinant human VISTA at pH 6.0 as previously described[37–39]. Positively selected sequences were reformatted as full-length human antibodies, produced via transient transfection of Expi293 cells, and validated for binding to VISTA by SPR. Where indicated, anti-human VISTA antibodies were converted to chimeric antibodies with a mouse IgG1-D265A isotype.

For imaging studies, VISTA.16 and VISTA.18 antibodies were fluorescently labelled with Alexa Fluor 680 using the SAIVI Rapid Antibody Labelling Kit (ThermoFisher). Monomericity of the labelled antibodies was confirmed by SEC. The amount of fluorophore conjugation was measured by LC–MS, and retention of VISTA binding was confirmed by SPR.

Anti-mouse VISTA antibodies were produced in VISTA-knockout mice immunized with recombinant mouse VISTA. Splenocytes from immunized mice were fused with the Sp2/0 myeloma cell line. Hybridoma supernatants were screened for reactivity to recombinant mouse VISTA by ELISA and to cell surface mouse VISTA by flow cytometry. Promising candidates were then subcloned, sequenced, and expressed recombinantly with a mouse IgG1-D265A isotype[40]. The clone VISTA.10, which bound mouse VISTA independent of pH and efficiently blocked its binding to T cells, was selected for use in mouse studies.

### Antibody epitope binning and pH sensitivity

Competitive SPR epitope binning was used to identify VISTA-specific antibodies which cross-block the desired VISTA.4 or undesired VISTA.5 epitopes using a Biacore T200 instrument. Antibodies VISTA.4 and VISTA.5 were diluted to 10 μg ml$^{-1}$ in 10 mM sodium acetate pH 4.5 and immobilized onto the flow cells of a CM5 biosensor following the manufacturer's amine coupling protocol (GE Healthcare). Competition was assessed at 25 °C using HBS-P+ running buffer (10 mM HEPES, 150 mM NaCl and 0.05% v/v Surfactant P20, pH 7.4). 100 nM monovalent human VISTA-ECD–His was captured by the immobilized antibodies, then each VISTA antibody screened was injected at 100 nM to evaluate VISTA.4 or VISTA.5-mediated co-binding or blocking activity. Two 30 s injections of 10 mM glycine pH 2.0 regenerated the VISTA.4 and VISTA.5 surfaces between assay cycles. Sensorgrams were analysed using Biacore T200 Evaluation Software v.2.0. Antibodies blocked by VISTA.4 and antibodies that were not blocked by VISTA.4 or VISTA.5 (that is, different epitope), were assessed for pH-dependent VISTA binding at pH 7.4 and 6.0.

### Human T cell functional assays

In co-culture experiments, CD4$^+$ T cells were enriched from healthy donor blood by negative selection (StemCell RosetteSep) and labelled with the proliferation dye CellTrace Violet (ThermoFisher). 293T cells were engineered to ectopically express a single chain variable fragment of the agonistic CD3 monoclonal antibody OKT3 and human VISTA (293T-OKT3-VISTA). CD4$^+$ T cells and irradiated 293T-OKT3-VISTA cells were co-cultured at a ratio of 4:1 in RPMI-1640 supplemented with 10% v/v heat-inactivated fetal calf serum (FCS), 2 mM L-glutamine (Gibco), 2 mM non-essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), 55 μM β-mercaptoethanol and titrated human anti-human VISTA or isotype-matched non-VISTA-binding control antibodies for 5 days. Proliferation was calculated as the percentage of CD4$^+$ T cells undergoing CellTrace Violet dye dilution, as determined by flow cytometry.

In NF-kB phosphorylation experiments, tissue culture-treated 96-well flat bottom plates were coated with OKT3 (0.5 μg ml$^{-1}$) and either human VISTA–Fc or isotype-matched control antibody (5.0 μg ml$^{-1}$) in PBS at 37 °C for approximately 2 h. Where indicated, VISTA–Fc proteins with H153, H154 and H155 residues mutated to alanine, aspartic acid, or arginine were used in place of wild-type VISTA–Fc. The wells were then washed with PBS and pre-incubated with human anti-human VISTA or non-VISTA-binding control antibodies diluted to 5.0 μg ml$^{-1}$ in HBSS acidified to various pH with MES for 30 min. T cells were suspended in the same HBSS + MES buffers, added to the wells, centrifuged, and cultured at 37 °C for 15 min. The cells were then fixed (Cytofix buffer, BD Biosciences), permeabilized (Phosflow Permeabilization Buffer III, BD Biosciences), and stained with anti-pNF-kB S529 (BD Biosciences) and anti-mIgG secondary antibodies (Jackson Immunoresearch) before being acquired on a flow cytometer. NF-kB phosphorylation was calculated as a percentage of the pNF-kB MFI for CD4+ T cells stimulated at pH 7.4 in wells that had been coated with OKT3 and isotype-matched control and pre-incubated with soluble non-binding isotype-matched control antibodies.

In Jurkat assays, Jurkat cells were engineered to express luciferase under the control of an Nf-KB-inducible promoter. These Jurkat NF-kB-luciferase cells were co-cultured with non-irradiated 293T-OKT3-VISTA cells at a ratio of 4:1 in HBSS (ThermoFisher) acidified to various pH with MES and human anti-human VISTA antibodies for 4 h. Jurkat cell activation was measured by luciferase substrate assay (Promega).

### Antibody epitope mapping

VISTA antibody binding epitopes were mapped by yeast display and NGS as previously described[41–43]. In brief, a saturation mutagenesis library of single point mutants of VISTA's extracellular domain was generated and displayed on the surface of yeast. VISTA mutants that lost binding to the test antibody, but retained binding to a non-cross-blocking VISTA antibody, were sorted and sequenced. The positions of the mutations in these mutants were designated as energetically important residues in the test antibody's epitope.

### SPR

VISTA antibody binding to recombinant human VISTA-ECD-His protein was measured at acidic and neutral pH (Biacore T200, GE Healthcare). Protein A (ThermoFisher catalogue no. 21181) was diluted to 20 μg ml$^{-1}$ in 10mM sodium acetate pH 4.5 and immobilized onto the flow cells of a CM5 biosensor following the manufacturer's amine coupling protocol (GE Healthcare). All SPR experiments were conducted at 37 °C using PBST (137 mM sodium chloride, 2.7 mM potassium chloride, 10 mM phosphate buffer and 0.05% Tween 20) running buffer at the indicated pH. Antibodies were diluted to 20 nM in PBST pH 7.4, and were captured on the protein A surface. A concentration series of 100–0.8 nM monovalent human VISTA-ECD–His was injected over the captured antibodies at 40 μl min$^{-1}$ to measure association and dissociation. Two 15 s injections of 10 mM glycine pH 1.5 regenerated the Protein A capture surface between assay cycles. Rate constants $k_a$ ($k_{on}$, association rate) and $k_d$ ($k_{off}$, disassociation rate) were derived from reference flow cell and 0 nM blank-subtracted sensorgrams, and were fit to a 1:1 binding model in Biacore T200 Evaluation Software v.3.1.

VISTA.18 Fab was prepared by papain digest following manufacturer's protocol (ThermoFisher catalogue no. 44985). Binding of VISTA.18 Fab was performed using the same running buffer (PBST) at pH 7.4 and 6.0. The previously described VISTA mutant Fc-fusion proteins were diluted to 25 nM and captured for 60 s at 10 μl min$^{-1}$ using an anti-human Fc CM5 sensor chip (GE Healthcare), following which VISTA.18 Fab was injected at concentration ranging from 0.8–100 nM. The surface was regenerated using a 30 s pulse of 3 M $MgCl_2$ at 30 μl min$^{-1}$. $\%R_{max}$ (normalized SPR binding response) for Fab binding was calculated as (observed $R_{max}$/calculated $R_{max}$) × 100, where observed $R_{max}$ = RU (binding response units) response at the end of the association phase; calculated $R_{max}$ = (capture level/molecular mass of ligand) × molecular mass of analyte × valency.

### Antibody engineering

Antibody variant libraries were built by introducing aspartate, glutamate and histidine substitutions in the CDRs of the heavy and light chain variable regions. The CDR variants were synthesized as oligo pools (Twist Biosciences), allowing for single and double amino acid substitutions in each CDR. Each library was constructed to allow for a maximum of six amino acid substitutions per chain. These libraries were expressed on the surface of yeast and subjected to several rounds of binding to recombinant human VISTA at pH 6.0. We then performed additional rounds of selection toggling between positive selection for VISTA binding at pH 6.0 and negative selection for VISTA binding at pH 7.4. Selected variants were reformatted as full-length human antibodies and validated by SPR.

### VISTA crystallography

N91Q, N108Q and N190Q mutations were introduced into VISTA's extracellular domain to reduce glycosylation. Recombinant VISTA.18 fragment antigen-binding (Fab) and low-glycosylation VISTA were labelled with AlexaFluor 488 and 555, respectively. The labelled proteins were co-incubated at a 4.8:1 molar ratio of VISTA:Fab in DPBS overnight at 4 °C. The resulting VISTA–Fab complex was purified by SEC (Superdex 200 16/200 gel filtration column, GE Healthcare). Fractions corresponding to the VISTA:Fab complex were pooled and concentrated to approximately 15 mg ml$^{-1}$. The complex was crystallized by combining 0.5 μl of the complex with 0.5 μl of the precipitant (1.8 M ammonium sulfate, 0.1 M phosphate/citrate, pH 4.2), over a reservoir containing 75 μl of additional precipitant, in an MRC UVXPO sitting drop vapour diffusion crystallization tray (Swissci) at room temperature. Initial crystals developed within 2 to 3 days, and grew to their full size of approximately 100 μM × 100 μM × 400 μM over approximately 10 days. Crystal fluorescence was visualized

using a RockImager 1000 crystal imaging system (Formulatrix). The crystals were cryoprotected by subermersion in 3.4 M ammonium sulfate and then flash frozen in liquid nitrogen for X-ray data collection.

X-ray diffraction data was collected at a wavelength of 1.0 Å and a temperature of 100 K at the Advanced Photon Source (IMCA-CAT beamline 17-ID). Data reduction was performed using HKL2000 (HKL Research). The VISTA:VISTA.18 Fab co-crystal structure was solved by molecular replacement (MR) using Phaser[44] and the coordinates from the VHVL and CH1CL portions of an internally determined Fab crystal structure as input models. The initial MR model provided enough phasing power to enable confident building of VISTA from scratch. The structure was completed through iterative cycles of model building using Coot[45] and restrained refinement using autoBUSTER (Global Phasing). A summary of the data collection and refinement statistics is provided in Supplementary Table 1. The VISTA:VISTA 18 Fab co-crystal structure has been deposited into the RCSB PDB under accession number 6MVL.

**Receptor-based ligand capture and mass spectrometry**

Receptor-based ligand capture and mass spectrometry was performed using TriCEPS (Dualsystems Biotech) as previously described[19], with the exception that the labelling buffer was acidified to pH 6.0. In brief, capture was performed on human $CD4^+$ T cells with VISTA–Fc chimeric protein bait. The anti-human CD3 antibody OKT3 was used as a control bait. For labelling, 300 μg each of OKT3 and human VISTA–Fc chimeric protein were buffer exchanged to 150 μl of 25 mM HEPES at pH 8.2. TriCEPS v. 3.03 (150 μg) was added to each reaction, mixed and incubated at room temperature with gentle shaking for 90 min. In parallel, 600 million human $CD4^+$ T cells were enriched from healthy donor blood by RosetteSep (StemCell) and suspended in PBS and 1% v/v heat-inactivated fetal calf serum at pH 6.0 (labelling buffer). The cells were then oxidized by treatment with 1.5 mM sodium metaperiodate at 4 °C for 15 min. After oxidation, the cells were washed, divided into two parts, and incubated with either the VISTA bait or the OKT3 bait at 4 °C with gentle shaking for 90 min. After labelling, the cells were washed, pelleted and snap frozen.

**VISTA, PSGL-1, P-selectin, KPL1, GP1BA, and VSIG-3 octet**

Binding interactions were detected using an OctetRed384 BLI instrument (PALL/ForteBio). All assay steps were conducted at 30 °C at 1,000 r.p.m. shake speed. Unless otherwise noted, the pH 6.0 buffer contained 50 mM MES, 200 mM sodium chloride, 4 mM calcium chloride and 0.05% v/v Tween 20. The pH 7.4 buffer contained 10 mM HEPES, 150 mM sodium chloride, 4 mM calcium chloride and 0.05% v/v Tween 20. These buffers were used for the full duration of the assays.

For binding experiments measuring VISTA, KPL1, and P-selectin binding to PSGL-1, the human PSGL-119-mer–Fc fusion protein (see recombinant proteins methods) was first captured on anti-human IgG–Fc sensors (AHC, PALL/ForteBio). Where specified, the human PSGL-119-mer–Fc fusion protein used was produced in cells not transfected with α1,3-fucosyltransferase and core 2 β1,6-*N*-acetylglucosaminyltransferase, or were separated by anion exchange liquid chromatography into sulfotyrosine-rich and sulfotyrosine-poor fractions. The anti-human capture sensors were then blocked with total human IgG (Jackson Immunoresearch). Binding to 500 nM wild-type human VISTA–Fc fusion protein (R&D Systems), 50 nM human P-selectin–Fc fusion protein (R&D Systems) and 200 nM KPL1 (R&D Systems) was measured for 10 min.

For antibody and VSIG-3 competition experiments, human VISTA–Fc was diluted to 400 nM and premixed for 30 min with 0 nM, 200 nM, 400 nM or 800 nM of each test antibody or human VSIG-3–Fc fusion protein (R&D Systems) before assessing binding. For KPL1 and human P-selectin competition experiments, captured human PSGL-119-mer–Fc fusion protein was blocked using 400 nM negative control antibody, KPL1 (Millipore Sigma) or 400 nM human P-selectin (R&D Systems), then dipped into titrated human VISTA–Fc.

For BLI binding experiments evaluating multi-pH interactions of VISTA to VSIG-3, CD42b/GP1Bα, PSGL-1 and VISTA, all assay steps were performed in DPBS buffer (Gibco) containing 0.05% v/v Tween 20, pH-adjusted to 5.8, 6.2, 6.6, 7.0 or 7.4 as indicated. The experimental conditions described above were applied, with the exception that 200 nM human VISTA–Fc was first captured to AHC sensors, and binding was instead measured to 500 nM wild-type human PSGL-119-mer–Fc fusion protein, 500 nM Y→A mutant human PSGL-119-mer–Fc fusion protein, 500 nM human VSIG-3–Fc fusion protein (R&D Systems), 500 nM human CD42b/GP1Bα (R&D Systems) or 500 nM human VISTA–Fc fusion protein (R&D Systems).

For BLI experiments measuring PSGL-1 binding to VISTA histidine mutants, VISTA H153/H154/H155 mutants with a human Fc tag at 500 nM were captured on anti-human IgG Fc sensor (Pall/Forte Bio) for 10 min, blocked with total human IgG (Jackson Immunoresearch) and dipped into human PSGL-119-mer–Fc (125, 250 and 500 nM) to measure association for 10 min.

Double reference-subtracted sensorgrams were analysed in Data Analysis 9.0 (PALL/ForteBio), and binding responses at the end of the association phase are reported.

**PSGL-1 glycopeptide ELISA**

Purified human PSGL-119mer–Fc proteins and control human Fc (Jackson ImmunoResearch) were adsorbed to ELISA plates (Nunc Maxisorb) at 10 μg $ml^{-1}$ in DPBS overnight at 4 °C. Plates were blocked with 1% BSA–0.05% Tween 20 (Teknova) for 1 h and dilutions of the anti-sLewisX monoclonal antibody HECA452 (Santa Cruz Biotechnology) made in the same buffer across the plate. Plates were incubated for 4 h at room temperature, washed and incubated for 1 h with anti-rat IgM-HRP (Jackson ImmunoResearch), followed by washing and detection with TMB substrate (Thermo). Plates were read in a Spectramax plus instrument using Softmaxpro (Molecular Devices).

For P-selectin ELISA assays, purified human PSGL-119mer–Fc proteins and control human Fc (Jackson ImmunoResearch) were adsorbed to ELISA plates (Nunc Maxisorb) at 1 μg $ml^{-1}$ in DPBS overnight at 4 °C. Plates were blocked with 1% BSA–0.05% Tween 20 (Teknova) with added 0.5 mM MgCl and 1 mM CaCl for 1 h. All subsequent additions were made in this same buffer. Dilutions of human P-selectin–Fc (R&D systems) were incubated 1 h at room temperature. After washing, biotinylated goat anti-Hu P-selectin (R&D systems) was added at 1:4,000 for 1 h at room temperature. This was followed by washing and adding streptavidin-HRP (Thermo) at 1:4000 for 1 h, then washing again and detection with TMB substrate (Thermo). Plates were read in a Spectramax plus instrument using Softmaxpro (Molecular Devices).

**Mass spectrometry**

PSGL-119-mer glycopeptide–Fc-fusion proteins were denatured in the presence of 0.5% Rapigest surfactant, reduced, alkylated, and digested by pepsin or trypsin and Glu-C. Data were acquired on QEPlus mass spectrometer (ThermoFisher) connected to Aquity UPLC (Waters) and analysed using Byonic software (Protein Metrics). Results were manually verified.

**Isothermal titration calorimetry**

VISTA–Fc, VISTA–His and PSGL-119-mer–Fc were dialysed against phosphate buffered saline (PBS) at pH 6.0 or pH 7.4 over 16 h at 4 °C. The concentrations of the proteins were determined by UV absorbance using Lunatic instrument (Unchained Labs). The concentrations used for the ITC experiments are listed in Supplementary Table 2. The ITC experiment with VISTA–Fc titration at 352 μM into 45 μM PSGL1–Fc at pH 7.4 was performed using a MicroCal PEAQ-ITC (Malvern Panalytical). This experiment involved a single injection of 0.4 μl followed by 18 injections of 2 μl with an injection duration of 4 s and a 150 s spacing between the injections. The reference power was set to 10 μcal $s^{-1}$, and the stirring speed was 750 r.p.m. Data processing was performed using MicroCal PEAQ-ITC analysis software v.1.21. The rest of the ITC experiments were performed by a MicroCal Auto-iTC200 (Malvern

Panalytical). These experiments started with a single injection of 0.5 µl followed by 19 injections of 2 µl with an injection duration of 4 s and an injection spacing of 180 s. The reference power and the stirring speed were 8 µcal $s^{-1}$ and 1,000 r.p.m., respectively. Data were processed using MicroCal Origin analysis software v.7.0 using the 'one set of sites' model. The fitted association constant ($K_A$) values were used to determine the dissociation constants ($K_D$) using the equation $K_D = 1/K_A$. The free energy of binding ($\Delta G$) was calculated using the relationship $\Delta G = -RT \ln K_A$ with $R$ and $T$ being the gas constant and the temperature, respectively.

### T cell CRISPR

Human $CD4^+$ T cells were enriched by RosetteSep (Miltenyi) from whole blood and stimulated with plate-coated antibodies against CD3 and CD28 (OKT3 and CD28.2 respectively, ThermoFisher). After 2 days, the cells were transfected in triplicate with Cas9 ribonuclear proteins (RNPs, Dharmacon) loaded with PSGL-1 guide RNAs (gRNAs) CACCAGCGCCAAGATTAGGA and CACTCAAAACCACAGCCATGG. gRNAs targeting CD4 and GFP (with no human homology) were used as controls. The transfected cells were then further stimulated with CD3/CD28 beads (Human T-activator Dynabeads, Miltenyi). After 4 days, VISTA multimer binding was assessed with gating on PSGL-1-knockout and PSGL-$1^+$ cells.

### PSGL-1–VISTA docking

The crystal structure of the human PSGL-1 19-mer peptide bound to human P-selectin contains coordinates for 13 of the 19 residues (QATEY-**EYLDYDFLPETEP**P, residues with coordinates in bold). The structure was modelled in Maestro v.10.7.015[46] by addition of the remaining six residues and energy minimization using the OPLS force field for peptides as implemented in the 'protein preparation' feature. The IgV domain of human VISTA was taken from the crystal structure reported here of anti-VISTA Fab complexed with VISTA. Missing side chains and loops were filled in using protein preparation in Maestro. The anti-VISTA Fab binding site was used for docking the PSGL-1 19-mer peptide to VISTA using the suggested protocol for protein–protein docking described in the BioLuminate module in Maestro.

### Wild-type and VISTA-knockout mice

Wild-type C57BL6/J mice were obtained from Jackson Laboratories and Charles River Laboratories. VISTA-knockout mice were obtained from the University of California Davis Knockout Mouse Project (KOMP).

### Human VISTA knock-in mice

A *Vista*-targeting vector was constructed from genomic C57BL/6N mouse strain DNA (genOway). A human *VISTA* sequence coding for the mature protein was inserted in frame with mouse exon 3 downstream of the mouse signal peptide, replacing the coding portion of exon 3 and part of intron 3. This insertion eliminates expression of the mouse *Vista* gene. A loxP-flanked neomycin resistance cassette was inserted in intron 2.

Linearized targeting vector was transfected with C57BL/6N embryonic stem cells. G-418 resistant embryonic stem cell clones were screened for locus recombination by PCR and Southern blot. The integrity of the humanized cassette was confirmed by sequencing.

Recombined embryonic stem cells were microinjected into C57BL/6N blastocysts, giving rise to male chimaeras. These mice were crossed with C57BL/6N mice expressing Cre recombinase to produce heterozygous *VISTA*-humanized mice devoid of the neomycin resistance cassette. The presence of the wild-type mouse allele (6.1 kb) and the humanized allele (7.1 kb) was confirmed by Southern blot using external probes of PciI digested DNA. Heterozygous animals were then interbred to produced homozygous *VISTA*-humanized mice.

### Fluorescence optical imaging

Mouse colorectal carcinoma MC38 cells were cultured in Dulbecco's Modified Eagle Media (DMEM) supplemented with 10% v/v heat-inactivated FCS in vitro. Six- to thirty-week-old female human VISTA knock-in mice were subcutaneously injected with $0.5–1.0 \times 10^6$ MC38 cells each while being fed an alfalfa-free diet (Teklad catalogue no. TD.97184) to minimize autofluorescence in the gastrointestinal tract. Tumour growth was monitored by caliper. When tumours reached an average volume of 70–110 $mm^3$, typically 7–10 days after implantation, mice were randomized into groups on the basis of tumour volume. Mice received a single intravenous injection of 3 mg $kg^{-1}$ Alexa Fluor 680-labelled VISTA.16 or VISTA.18. Subsets of mice were euthanized at 2.5, 24 and 51 h post-injection for imaging of heart, liver, spleen, kidney, lung, stomach, intestine, muscle, tumour and whole blood. Images were acquired under the following parameters: excitation = 679 nm, emission = 702 nm, exposure = auto-setting, binning = 8, F/stop = 2, FOV = 22.3 cm and focus = 1.5 cm. Fluorescence intensities were quantified in units of radiant efficiency using Living Image software (v.4.5, PerkinElmer). The average intensities of the tissues of interest were normalized to the average number of fluorophores conjugated to VISTA.16 (fluorophore to antibody ratio, 0.79) and VISTA.18 (fluorophore to antibody ratio, 1.31) and set to the same colour scale before defining and representing regions of interest.

### Tumour studies

Mouse colorectal carcinoma MC38 cells were cultured as described above. Six to twelve-week-old female C57BL6/J mice, VISTA-knockout mice and their wild-type littermates, and human VISTA-knock-in mice were subcutaneously injected with $0.5–1.0 \times 10^6$ MC38 cells per mouse. Tumour growth was monitored by caliper. When tumours reached an average volume of 70–110 $mm^3$, typically 7–10 days after implantation, mice were randomized into groups on the basis of tumour volume. Mice received 100 µg anti-PD-1, 600 µg anti-VISTA (mouse- or human-reactive as indicated), and/or 600 µg of a control anti-diptheria toxin IgG1-D265A antibody administered four times by intraperitoneal injection every 3–4 days. All antibodies were IgG1-D265A (Fc-inert) isotype. Researchers were not blinded to treatment assignment. Animals were continuously monitored, and mice were euthanized via asphyxiation when any of the following endpoints were met: study termination, tumour burden equal or greater than 2,000 $mm^3$, tumour ulceration, body weight loss equal or greater than 20%, or moribund appearance. Mice whose tumours were unmeasurable or below 10 $mm^3$ were considered to be tumour-free.

### Ex vivo mouse leukocyte profiling

Where indicated, the tumours, spleens, and/or blood of mice were profiled ex vivo 7 days after the start of treatment. Tumours were enzymatically disassociated with 250 U $ml^{-1}$ collagenase IV (ThermoFisher) and 100 mg $ml^{-1}$ DNase I (Roche) in HBSS supplemented with 5% v/v heat-inactivated FCS and 5 mM $CaCl_2$ and mechanically disassociated with GentleMacs cell disruptors (Miltenyi). Spleens were mechanically disassociated only. Spleen and blood samples were treated with ACK red blood cell lysis buffer (ThermoFisher). Cells were stained with the antibodies CD3, CD4, CD8, CD11b, CD45, F4/80, FoxP3, Ly6C, Gr1, MHC-II, mouse VISTA, human VISTA, PD-1, LAG-3 and TIM-3. Blocking was performed with anti-CD16/32 (2.4G2, BD Biosciences). Intracellular staining was performed with the FoxP3 Fixation/Permeabilization buffer kit (ThermoFisher).

### Antibody pharmacokinetics

In mice, VISTA.16 and VISTA.18 were evaluated following intravenous injections into 6–12 week old female C57BL6/J mice as well as human VISTA knock-in mice and their wild-type littermates at a dose of 5 mg $kg^{-1}$ ($n$ = 1–4 mice per antibody and genotype as indicated). Serial blood samples were collected at 0.25, 1, 6, 24, 48, 96, 168, 264, 336 and 504 h post-injection.

In cynomolgus macaques, VISTA.4 and VISTA.18 were evaluated following 10 min intravenous infusions into protein-naive 5–6 year old male cynomolgus monkeys at a dose of 5 mg $kg^{-1}$ ($n$ = 1 per antibody). Serial blood samples were collected at 0.17, 0.5, 2, 4, 6, 24, 48, 72, 168, 216, 240 and 336 h post-infusion.

Subsequently, serum samples were obtained for antibody concentration analysis using a ligand-binding assay that employed the recombinant VISTA as a capturing agent and an anti-human IgG Fc mAb as a detecting agent. The lower limit of quantification for the assay was 1 ng $ml^{-1}$. Mean residence times were estimated by non-compartment analysis of the serum mAb concentration-time data using Kinetica software (v.5.0, Thermo Fisher Scientific).

### Statistics

Except where indicated, statistics depict means, standard errors of the mean, and one-way ANOVA with Dunnett's multiple comparisons. Receptor-based ligand capture statistics were performed as previously described[47].

### Research ethics

Human blood was obtained from a research blood donation program administered by the Bristol-Myers Squibb Occupational Health and Wellness department. The program was operated in compliance with all relevant ethical regulations, and written informed consent was obtained from all donors.

Animal studies were conducted in compliance with all relevant ethical regulations. Animal studies performed at Bristol-Myers Squibb were approved by the Bristol-Myers Squibb Institutional Care and Animal Use Committee. Animal studies performed at Five Prime Therapeutics were approved by the Five Prime Therapeutics Institutional Care and Animal Use Committee. Animal studies performed at Murigenics were approved by the Bristol-Myers Squibb Animal Welfare Risk Assessment Team and by the Murigenics Institutional Care and Animal Use Committee.

### Materials availability

All unique biological materials are available subject to a material transfer agreement.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

All data are available from the corresponding author and have been included in the manuscript or Supplementary Information. The VISTA:VISTA.18 Fab co-crystal structure has been deposited into the Protein Data Bank under accession number 6MVL.


36. Lonberg, N. Human antibodies from transgenic animals. *Nat. Biotechnol.* **23**, 1117–1125 (2005).
37. Roberts, R. W. & Szostak, J. W. RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proc. Natl Acad. Sci. USA* **94**, 12297–12302 (1997).
38. Kurz, M., Gu, K. & Lohse, P. A. Psoralen photo-crosslinked mRNA–puromycin conjugates: a novel template for the rapid and facile preparation of mRNA–protein fusions. *Nucleic Acids Res.* **28**, E83 (2000).
39. Xu, L. et al. Directed evolution of high-affinity antibody mimics using mRNA display. *Chem. Biol.* **9**, 933–942 (2002).
40. Clynes, R. A., Towers, T. L., Presta, L. G. & Ravetch, J. V. Inhibitory Fc receptors modulate in vivo cytotoxicity against tumor targets. *Nat. Med.* **6**, 443–446 (2000).
41. Chao, G., Cochran, J. R. & Wittrup, K. D. Fine epitope mapping of anti-epidermal growth factor receptor antibodies through random mutagenesis and yeast surface display. *J. Mol. Biol.* **342**, 539–550 (2004).
42. Oliphant, T. et al. Antibody recognition and neutralization determinants on domains I and II of West Nile Virus envelope protein. *J. Virol.* **80**, 12149–12159 (2006).
43. Kowalsky, C. A. et al. Rapid fine conformational epitope mapping using comprehensive mutagenesis and deep sequencing. *J. Biol. Chem.* **290**, 26457–26470 (2015).
44. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
45. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
46. Maestro (Schrödinger, 2019).
47. Frei, A. P., Moest, H., Novy, K. & Wollscheid, B. Ligand-based receptor identification on living cells and tissues using TRICEPS. *Nat. Protoc.* **8**, 1321–1336 (2013).

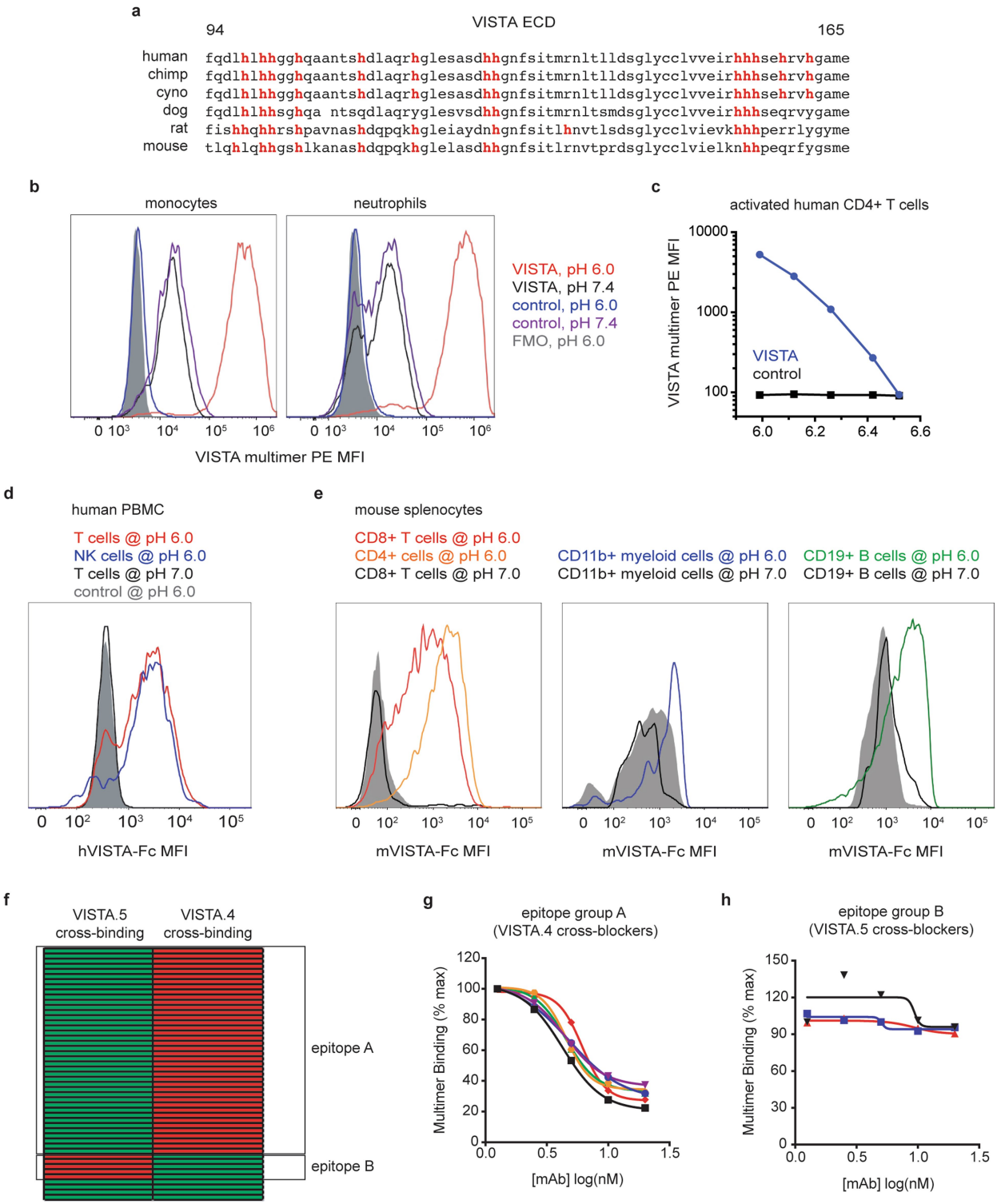


**Acknowledgements** We thank D. Ardourel, C. Bandoski-Gralinski, C. Bee, G. Bolton, M. Broz, I. Chakraborty, C. Connelly, F. Denhez, C. Gao, L. Garrenton, J. Gordon, N. Hammond, S. Hatcher, C. Hollander, M. Han, M. Happer, P. Helbling, V. Jenny, M. Labrecque, D. Myszka, A. Nallakkan, M. A. Pazos, P. Isnard-Petit, R. Lan, T. Metzger, E. Musteata, B. Nichols, M. Rodriguez, F. Sonego, S. Trouttet-Masson, S. Santino, E. Seo, T. Sproul, J. Sung, M. Supe, H. Tang, C. Terragni, J. Toth, S. Walrond and D. Wensel for technical assistance. We thank F. Bahjat, R. Camphausen, G. Cantor, G Chen, P. Chow, R. Das Gupta, A. Dongre, P. Haroldsen, M.-C. Gaudreau, D. Lipovsek, N. Lonberg, J. Muckelbauer, S. Mueller, X. M. Schebye, M. Selby, P. Strop, D. Tenney and M. Wright for discussions. We thank C. Bolger for editorial support.


**Author contributions** R.J.J., L.J.S., M.Q., A.R., K.S.B. and A.J.K. conceived and supervised the study. R.J.J., A.Krishnakumar, E.B., A.L.R., H.L., T.C., H.F. and Y.-K.W. conducted biology experiments. L.J.S., J.P., R.D., L.C., G.R. and H.C. conducted antibody campaigns. J.P., X.D., M.C., R.Y.-C.H., B.E., J.M.R., A.P.Y. and S.D. conducted biophysical experiments. D.C., A.N., X.D., A.R., P.S. and J.H. conducted crystallography experiments and structural analyses. G.H.M. and K.T. generated the human VISTA knock-in mouse. J.N. and A.O.S. conducted imaging experiments. Z.Y., R.R. and A.Kozhich conducted pharmacokinetic experiments. R.J.J. and A.J.K. wrote the manuscript with input from other authors.

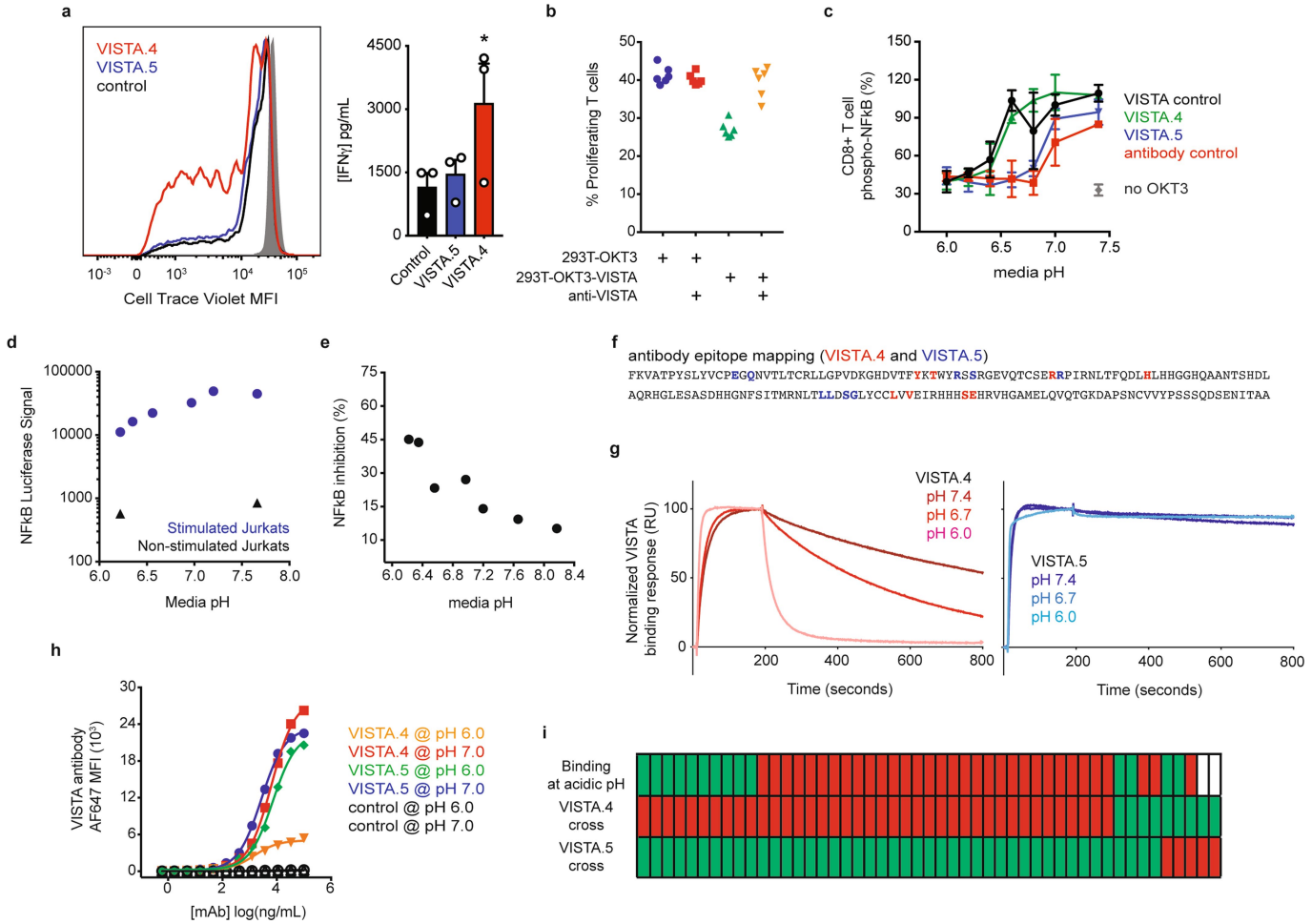
**Competing interests** All authors are or were employees of the companies Bristol-Myers Squibb, Five Prime Therapeutics, and genOway, which develop drugs and research models for profit.


**Additional information**
**Supplementary information** is available for this paper at https://doi.org/10.1038/s41586-019-1674-5.
**Correspondence and requests for materials** should be addressed to R.J.J.
**Peer review information** *Nature* thanks Linda M. Bradley, Gordon Freeman, Christopher Garcia and the other anonymous reviewer(s) for their contribution to the peer review of this work.
**Reprints and permissions information** is available at http://www.nature.com/reprints.

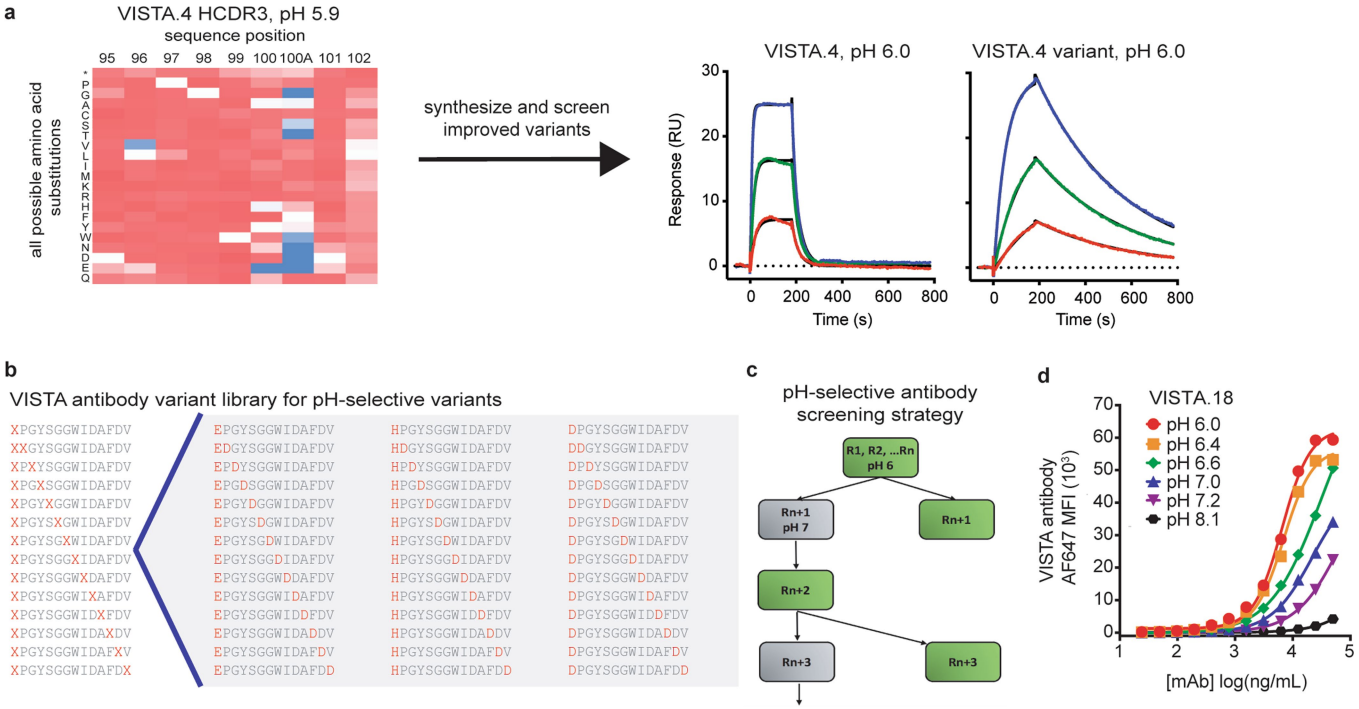



**Extended Data Fig. 1** | See next page for caption.

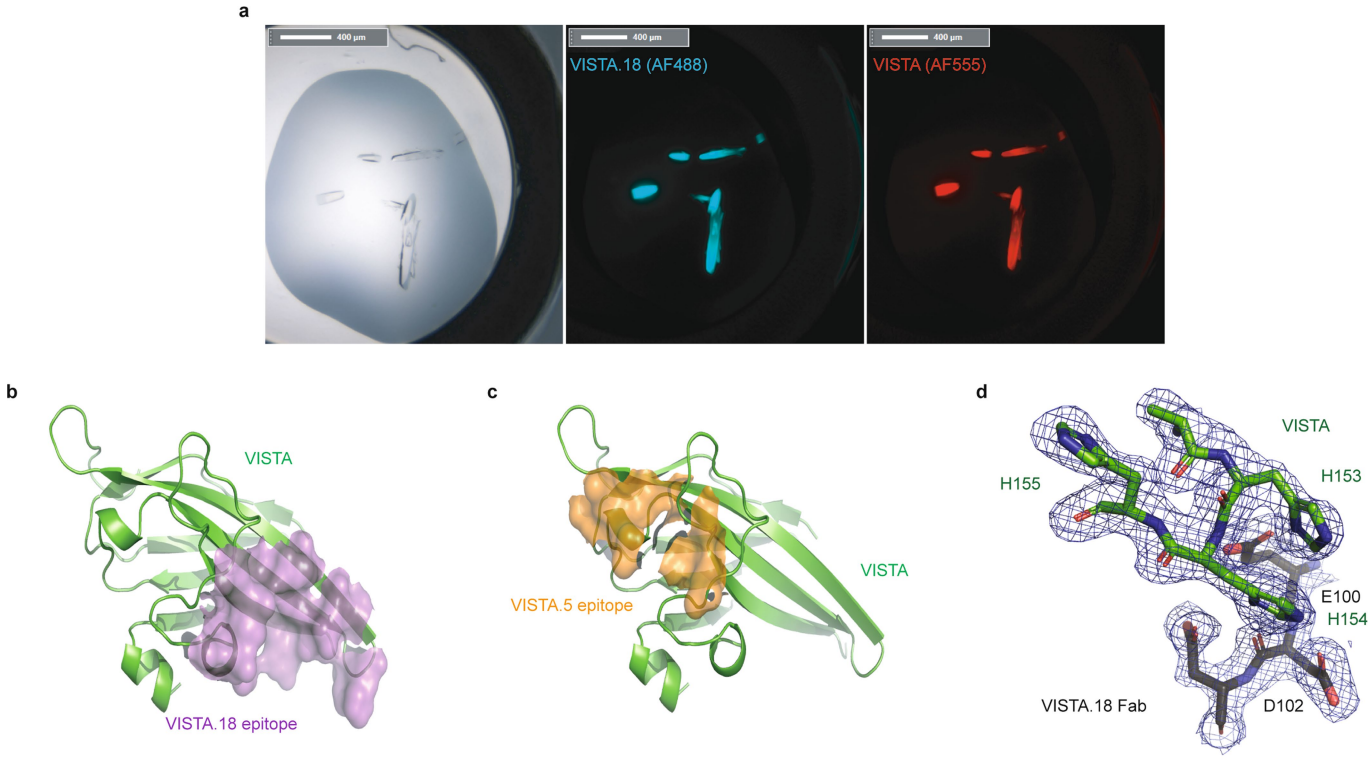
**Extended Data Fig. 1 | Conservation of VISTA pH selectivity. a**, Alignment of human VISTA extracellular domain amino acid residues 94–165 with chimpanzee, cynomolgus macaque, dog, rat and mouse equivalents. Histidine residues are highlighted in red. **b**, Human monocytes (left) and neutrophils (right) labelled with VISTA multimers at pH 6.0 and pH 7.4. Cells labelled with non-VISTA-loaded multimers (control) or left unstained (FMO) are also depicted. These data are representative of two independent experiments. **c**, VISTA multimer (blue) and non-VISTA-loaded multimer (control, black) binding to the activated human T cells depicted in Fig. 1c. Data are VISTA multimer MFI and are representative of six independent experiments. **d**, Human VISTA–Fc binding to human PBMC NK cells at pH 6.0 (blue), T cells at pH 6.0 (red) and T cells at pH 7.0 (black). T cells stained at pH 6.0 with the anti-human Fc secondary but not with VISTA–Fc are included as a control (grey filled). These data are representative of ten independent experiments. **e**, Mouse VISTA–Fc binding at pH 6.0 to wild-type mouse splenic $CD8^+$ T cells (red), $CD4^+$ T cells (orange), B cells (green) and $CD11b^+$ myeloid cells (blue). VISTA–Fc binding at pH 7.0 (black) and isotype-matched human IgG binding at pH 6.0 (grey filled) are included as controls. These data are representative of five independent experiments. **f**, Competitive SPR epitope binning of VISTA-specific antibodies against VISTA.4 and VISTA.5. Each row represents a unique clone, and for each clone, green indicates no cross-blocking, and red indicates cross-blocking. These data are representative of one experiment. **g**, **h**, Antibody blocking of VISTA-multimer binding to T cells as described in Fig. 1. Data are VISTA multimer MFI normalized to control and are representative of one experiment. **g**, Blocking activity by antibodies from the VISTA.4 epitope bin. VISTA.4 itself is depicted as black squares. **h**, Blocking activity by antibodies from the VISTA.5 epitope bin. VISTA.5 itself is depicted as black downwards triangles.




**Extended Data Fig. 2 | Effects of pH on VISTA function and antibody binding.** **a**, Representative histograms of CellTrace Violet dilution (left) and supernatant IFN-γ (right) by $CD4^+$ T cells co-cultured with 293T-OKT3-VISTA cells in the presence in VISTA.4 (red), VISTA.5 (blue), a non-VISTA-binding isotype-matched antibody (control, black), or without 293T-OKT3-VISTA cells (grey filled). Data are mean ± s.e.m. with one-way ANOVA and Dunnett's multiple comparisons. *$P = 0.0498$. $n = 3$ T cell donors; these data are representative of seven independent experiments. **b**, Per cent of $CD4^+$ T cells that proliferated following co-culture with 293T-OKT3-VISTA or 293T-OKT3 cells and VISTA.4 or an isotype-matched non-VISTA-binding control antibody. These data are representative of two independent experiments. **c**, NFκB phosphorylation in human $CD4^+$ T cells following stimulation with plate-coated OKT3 and VISTA–Fc in the presence of the antibodies VISTA.4 (green upward triangles), VISTA.5 (blue downward triangles) and a non-VISTA-binding control (antibody control, red squares) at various pH. Cells stimulated with OKT3 and a plate-coated control antibody (VISTA control, black circles) and without OKT3 (grey diamond) are included as controls. Data are mean ± s.e.m. pNF-κB MFI normalized to control. $n = 2$ T cell donors; these data are representative of two independent experiments. **d**, **e**, Jurkat NFκB-luciferase reporter cells were co-cultured with 293T-OKT3 or 293T-OKT3-VISTA cells and with VISTA.4 or non-VISTA-binding isotype-matched control antibody. These data are a composite of three independent experiments. **d**, Luciferase signal after culture with 293T-OKT3 cells (blue circles) or without 293T cells (black triangles). **e**, Per cent increase in the luciferase signal with VISTA.4 treatment during culture with 293T-OKT3-VISTA cells. **f**, VISTA.4 (red) and VISTA.5 (blue) binding epitopes on the human VISTA extracellular domain. **g**, Human VISTA SPR binding sensorgrams for the blocking antibody VISTA.4 (left; pH 6.0, light red; pH 6.7, red; pH 7.4, dark red) and the non-blocking antibody VISTA.5 (right; pH 6.0, light blue; pH 6.7, blue; pH 7.4, dark blue). Overlaid sensorgrams are 100 nM VISTA binding responses, normalized to the binding report point. These data are representative of six independent experiments. **h**, Cell binding of VISTA.4 (pH 6.0, orange downward triangles; pH 7.0, red squares), VISTA.5 (pH 6.0, green diamonds; pH 7.0, blue circles), and a non-VISTA-binding antibody (pH 6.0, unfilled circles; pH 7.0, unfilled upward triangles) to Raji cells ectopically expressing VISTA. Data are VISTA antibody MFI and are representative of five independent experiments. **i**, VISTA antibody epitope binning against VISTA.4 (centre row) and VISTA.5 (bottom row). Each row represents a unique clone, and for each clone, green indicates a lack of cross-blocking and red indicates cross-blocking. Binding capacity at pH 6.0 relative to binding capacity at pH 7.4 is also depicted (top row). For binding at acidic pH, red indicates a greater than threefold impairment in $k_d$ at pH 6.0, green indicates a less than threefold impairment, and white indicates no data. These data are representative of one experiment.



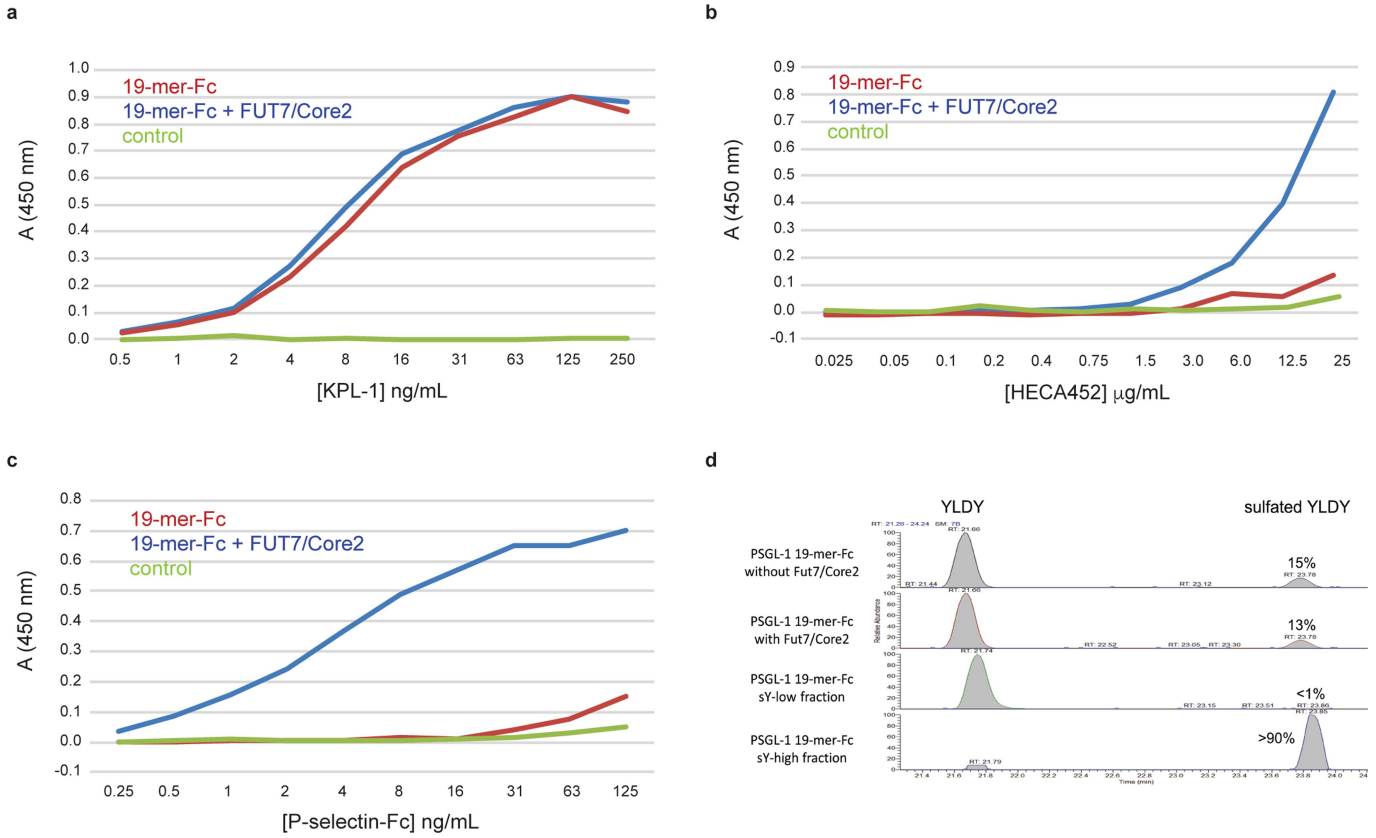

**Extended Data Fig. 3 | Acidic pH-selective antibody engineering.** **a**, Schematic depicting the method by which the VISTA.4 antibody was engineered to identify variants with improved binding at acidic pH. **b**, Schematic depicting the libraries of VISTA antibody variants used for screening acidic pH-selective variants. **c**, Schematic depicting the iterative screening strategy for identification of acidic pH-selective VISTA antibody variants. **d**, Cell binding of the acidic pH-selective antibody VISTA.18 to Raji cells ectopically expressing VISTA at pH 6.0 (red circles), pH 6.4 (orange squares), pH 6.6 (green diamonds), pH 7.0 (blue upward triangles), pH 7.2 (purple downward triangles) and pH 8.1 (black hexagons). Data are VISTA antibody MFI and are representative of three independent experiments.



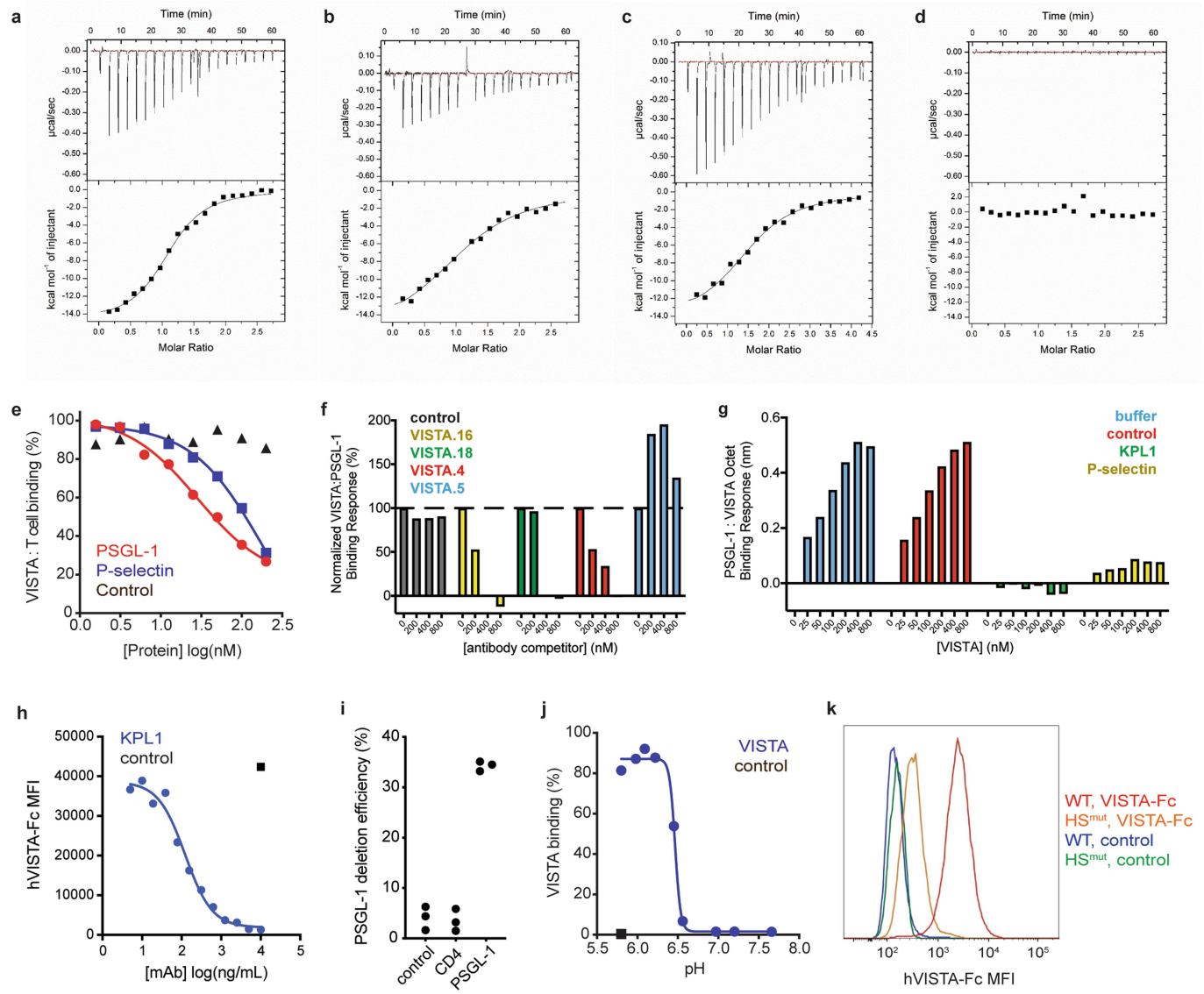

e

| Variant | 7.4 kd (1/s) | 7.4 KD (M) | 7.4 %Rmax | 6.7 kd (1/s) | 6.7 KD (M) | 6.7 %Rmax | 6.0 kd (1/s) | 6.0 KD (M) | 6.0 %Rmax |
|---|---|---|---|---|---|---|---|---|---|
| VISTA.18 D52N_D102V | 6.7E-03 | 8.7E-08 | 60.0 | 1.8E-03 | 7.5E-09 | 101.0 | 1.2E-03 | 8.8E-10 | 114.6 |
| VISTA.18 D52N | | | 3.9 | 6.7E-03 | 2.1E-07 | 38.9 | 1.5E-03 | 3.2E-09 | 116.6 |
| VISTA.18 D52N_E55A | | | 2.3 | 5.5E-03 | 1.2E-07 | 33.7 | 1.5E-03 | 4.0E-09 | 112.0 |
| VISTA.18 E55A_D102V | 1.0E-02 | 2.8E-07 | 35.4 | 3.1E-03 | 1.5E-08 | 85.7 | 2.0E-03 | 1.9E-09 | 111.3 |
| VISTA.18 D102V | 9.2E-03 | 1.7E-07 | 44.7 | 2.8E-03 | 1.4E-08 | 98.5 | 2.0E-03 | 1.3E-09 | 111.9 |
| VISTA.18 E55A | | | 2.8 | 8.5E-03 | 3.8E-07 | 26.8 | 2.3E-03 | 5.7E-09 | 108.0 |
| VISTA.18 E30D_D52N | | | 4.5 | 1.1E-02 | 1.2E-07 | 35.8 | 2.3E-03 | 5.8E-09 | 105.7 |
| VISTA.18 | | | 1.9 | 7.2E-03 | 1.8E-07 | 31.1 | 2.4E-03 | 5.2E-09 | 111.7 |
| VISTA.18 E30D_D102V | 1.1E-02 | 2.2E-07 | 40.8 | 3.5E-03 | 1.7E-08 | 95.7 | 3.0E-03 | 2.1E-09 | 114.8 |
| VISTA.18 E30D | | | 5.8 | | | 21.0 | 3.2E-03 | 9.6E-09 | 99.8 |
| VISTA.18 E30D_E55A | | | 3.5 | | | 22.6 | 3.5E-03 | 8.3E-09 | 108.8 |
| VISTA.16 (VISTA.18 parent) | 7.5E-03 | 3.2E-08 | 93.4 | 4.4E-03 | 1.1E-08 | 96.9 | 8.2E-03 | 4.5E-09 | 98.5 |
| VISTA.18 E100fF_D102V | 8.3E-03 | 4.1E-08 | 86.8 | 4.7E-03 | 1.1E-08 | 103.4 | 8.2E-03 | 2.7E-09 | 103.4 |
| VISTA.18 E55A_E100fF | | | 1.8 | | | 8.6 | 1.3E-02 | 6.2E-08 | 42.5 |
| VISTA.18 D52N_E100fF | | | 9.4 | 5.1E+01 | 3.9E-08 | 22.7 | 6.6E-02 | 2.0E-08 | 64.0 |
| VISTA.18 E100fF | | | 1.8 | 2.6E-02 | 5.2E-06 | 9.6 | 2.8E+00 | 4.2E-08 | 49.4 |
| VISTA.18 E30D_E100fF | | | 1.0 | 2.7E-02 | 1.0E-05 | 7.0 | 6.3E+01 | 8.2E-08 | 32.9 |
| control | | | -0.4 | | | 0.1 | | | -0.4 |

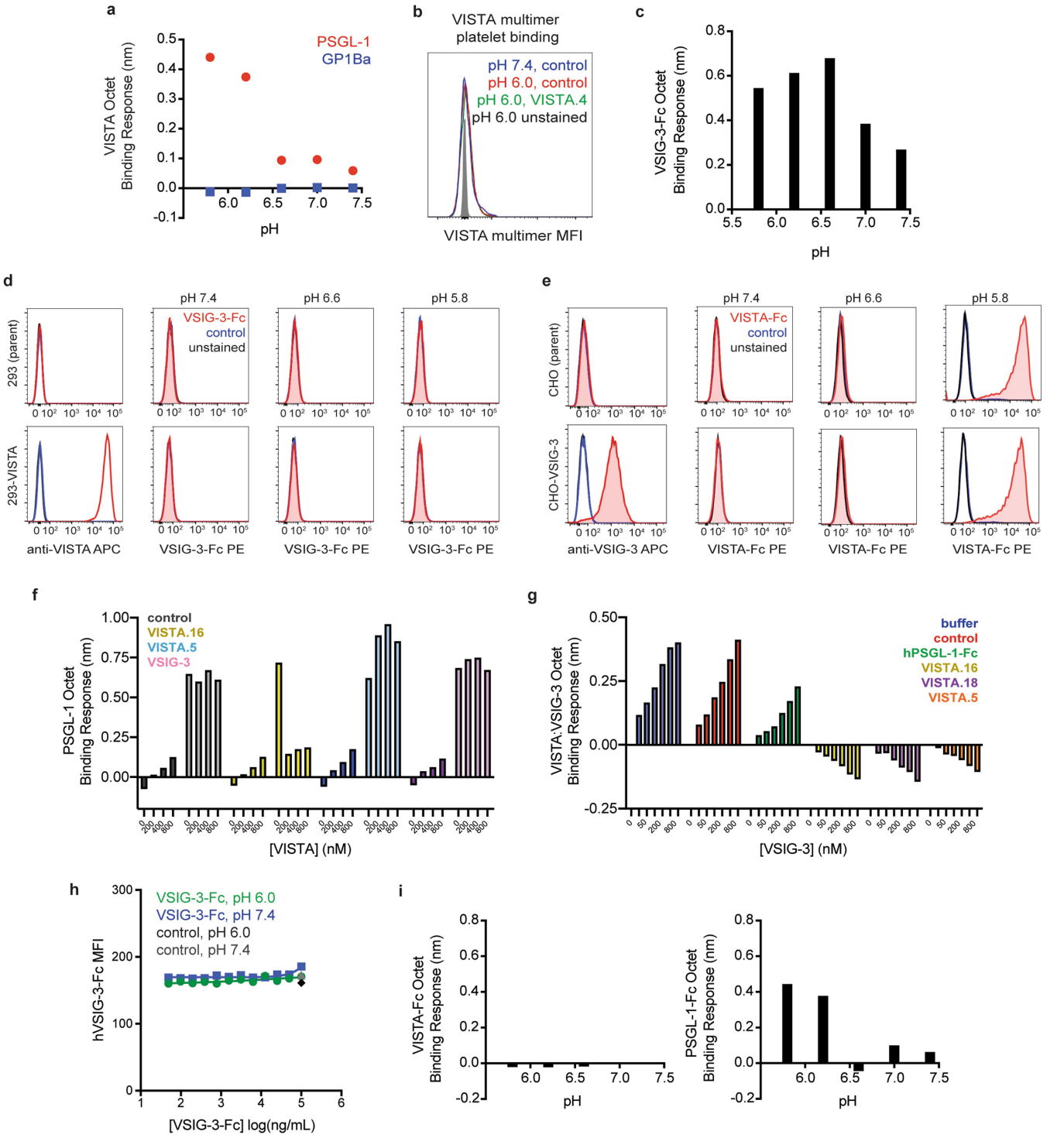
**Extended Data Fig. 4 | Co-crystallization of VISTA and VISTA.18.** The VISTA IgV domain (labelled with Alexa Fluor 555) and the VISTA.18 Fab (labelled with Alexa Fluor 488) were co-crystallized as described in Fig. 4 and the Methods. **a**, Representative bright-field (left), Alexa Fluor 488 fluorescence (centre), and Alexa Fluor 555 fluorescence (right) images of the crystals, indicating the presence of both VISTA and VISTA.18. These data are representative of one experiment. **b**, **c**, Superimpositions of the molecular surfaces of the yeast display-defined epitopes for VISTA.18 (purple, **b**) and VISTA.5 (non-blocking, orange, **c**) on the VISTA IgV domain (green). **d**, 2m|$F_O$ – D$F_C$| electron density map (blue mesh) contoured to 1$\sigma$ about the VISTA histidine triad (green sticks) and VISTA.18 HCDR3 (grey sticks). **e**, Human VISTA SPR binding data for the acidic pH-selective antibody VISTA.18 and variants of VISTA.18 in which the indicated residues have been reverted back to their identity in VISTA.18's non-acidic-pH-selective parent, VISTA.16. VISTA.16 and a non-VISTA binding isotype-matched control antibody are included as controls. These data are representative of one experiment.




**Extended Data Fig. 5 | PSGL-1 glycopeptide characterization. a**–**c**, ELISA binding curves of human PSGL-1 19-mer–Fc proteins produced with (blue lines) or without (red lines) FUT7 and Core2 co-transfection. Binding curves for isotype-matched control IgG are also shown (green lines). Data are absorbance at 450 nm and are representative of three independent experiments. **a**, Binding to the anti-human PSGL-1 antibody KPL1. **b**, Binding to the anti-sialyl-Lewis X moiety antibody HECA452. **c**, Binding to recombinant human P-selectin–Fc. **d**, Extracted ion chromatograms of the peptide YLDY in PSGL-1 19-mer–Fc proteins produced with or without FUT7 and Core2 co-transfection and with or without fractionation as indicated. The percentage of total YLDY that was sulfated is indicated for each sample. These data are representative of one experiment.

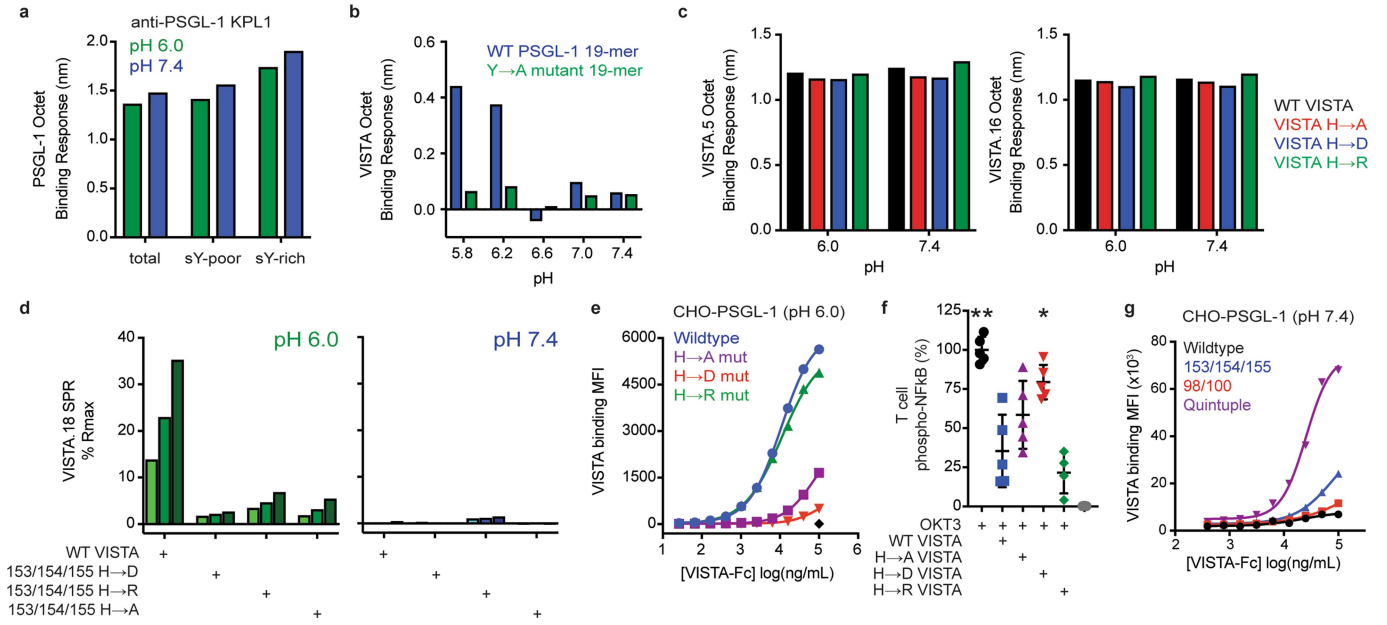



**Extended Data Fig. 6 | Further characterization of VISTA binding to PSGL-1.** **a**–**d**, Isothermal titration calorimetry (ITC) measurements of the interaction between PSGL-1 and VISTA. Top plots depict the raw calorimetric data of the titrations, and the bottom plots depict the integrated data corrected for the heat of dilution. The one set of sites model was used for data fitting. These data are representative of one experiment. **a**, Titration of 130 µM PSGL-1–Fc into 10 µM VISTA–Fc at 25 °C and pH 6.0. **b**, Titration of 130 µM PSGL-1–Fc into 10 µM VISTA–Fc at 37 °C and pH 6.0. **c**, Titration of 200 µM PSGL-1–Fc into 10 µM VISTA–His at 25 °C and pH 6.0. **d**, Titration of 130 µM PSGL-1–Fc into 10 µM VISTA–Fc at 25 °C and pH 7.4. The thermodynamic parameters determined by ITC are listed in Supplementary Table 2. **e**, Effects of PSGL-1–Fc (red circles) and P-selectin–Fc (blue squares) recombinant proteins on VISTA multimer binding to activated human $CD4^+$ T cells. A non-binding antibody (black triangles) is included as a control. Data are VISTA multimer MFI normalized to control and are representative of two independent experiments. **f**, Effects of the indicated VISTA antibodies on PSGL-119-mer–Fc fusion protein Octet binding to VISTA–Fc fusion protein at pH 6.0. A non-VISTA-binding antibody is included as a control. Data are BLI binding magnitudes and are representative of two independent experiments. **g**, Effects of the PSGL-1 antibody KPL1 and recombinant P-selectin on human PSGL-119-mer–Fc fusion protein Octet binding to VISTA–Fc fusion protein at pH 6.0. A non-PSGL-1-binding antibody (control) and no added antibody are included as controls. Data are BLI binding magnitudes and are representative of two independent experiments. **h**, Effects of KPL1 (blue circles) on human VISTA–Fc binding to human PBMC monocytes at pH 6.0. An isotype-matched non-PSGL-1-binding antibody is included as a control. Data are VISTA-Fc MFI and are representative of two independent experiments. **i**, Percentage of T cells with no PSGL-1 expression after CRISPR using guides against PSGL-1, CD4 or a scrambled control. These data are representative of five independent experiments. **j**, VISTA multimer (blue circle) binding to CHO cells expressing PSGL-1 at various pH. Non-VISTA-loaded multimer (black square) binding is included as a control. Data are VISTA multimer MFI normalized to control and are representative of two independent experiments. **k**, Human VISTA–Fc binding to wild-type and heparan sulfate-deficient CHO-K1 cells (red and orange, respectively) at pH 6.0. Isotype-matched control antibody binding to wild-type and heparan sulfate-deficient CHO-K1 cells (blue and green, respectively) at pH 6.0 is also shown. These data are representative of two independent experiments.

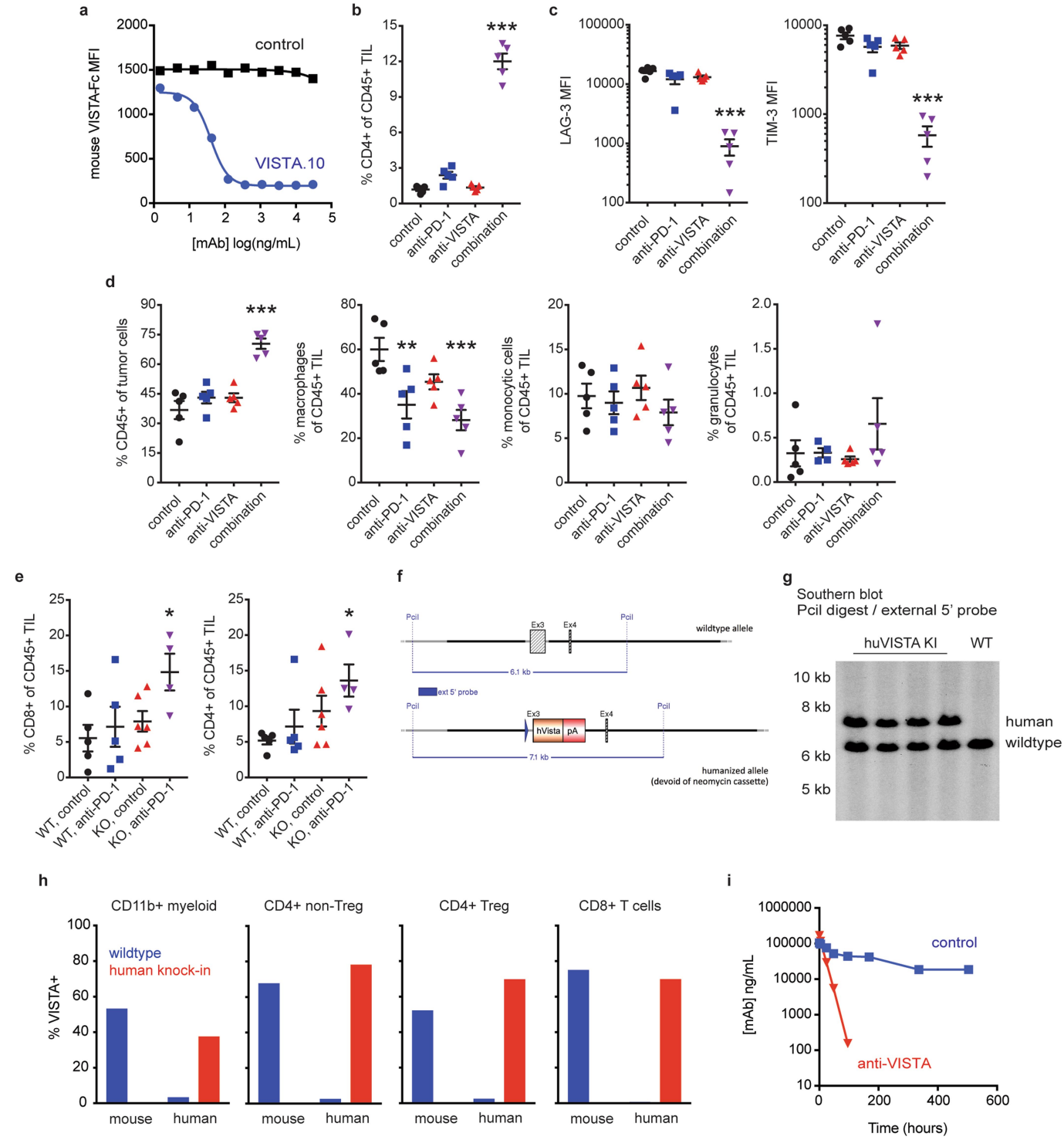



**Extended Data Fig. 7** | See next page for caption.

**Extended Data Fig. 7 | Other candidate VISTA receptors. a**, BLI binding magnitudes for GP1BA–His (blue squares) and PSGL-119-mer–Fc (red circles) to captured VISTA–Fc at the indicated pH. These data are representative of one experiment. **b**, VISTA multimer binding histograms to human platelets. Binding was performed in the presence of non-VISTA-binding control antibodies (pH 7.4, blue; pH 6.0, red) or the VISTA.4 blocking antibody (pH 6.0, green). Unstained platelets (grey filled histogram) are included as controls. These data are representative of two independent experiments. **c**, BLI binding magnitudes for VSIG-3–Fc binding to captured VISTA–Fc at the indicated pH. These data are representative of two independent experiments. **d**, Left, anti-VISTA stained (red), control stained (blue) or unstained (black) parental HEK293 (top) and VISTA-expressing HEK293 cells (bottom). Right, VSIG-3–Fc (red) or control–Fc (blue) binding to the same cells at the indicated pH. These data are representative of two independent experiments. **e**, Left, anti-VSIG-3 stained (red), control stained (blue) or unstained (black) parental CHO (top) and VSIG-3-expressing CHO cells (bottom). Right, VISTA–Fc (red) or control–Fc (blue) binding to the same cells at the indicated pH. These data are representative of two independent experiments. **f**, BLI binding magnitudes of VISTA–Fc binding to captured PSGL-119-mer–Fc at pH 6.0. Competition was provided at the indicated concentrations by a non-binding control antibody (grey bars), the VISTA blocking antibody VISTA.16 (yellow bars), the VISTA non-blocking antibody VISTA.5 (blue bars) or human VSIG-3–Fc fusion protein (purple bars). Darker bars depict the BLI binding magnitudes of competitors without VISTA. These data are representative of one experiment. **g**, BLI binding magnitudes of VSIG-3–Fc at the indicated concentrations binding to captured VISTA–Fc at pH 6.0. Competition was provided by buffer alone (blue bars), human PSGL-119-mer–Fc (green bars), non-binding isotype matched control antibody (red bars), VISTA.16 (yellow bars), VISTA.18 (purple bars), or VISTA.5 (orange bars). These data are representative of one experiment. **h**, VSIG-3–Fc binding to activated human PBMC T cells at pH 6.0 (green circles) or pH 7.4 (blue squares). Binding of isotype-matched control antibody at pH 6.0 (black diamond) and pH 7.4 (grey triangle) is included as a control. Data are VSIG-3–Fc MFI and are representative of two independent experiments. **i**, BLI binding magnitudes for VISTA–Fc (left) and PSGL-119-mer–Fc (right) binding to captured VISTA–Fc at the indicated pH. These data are representative of one experiment.

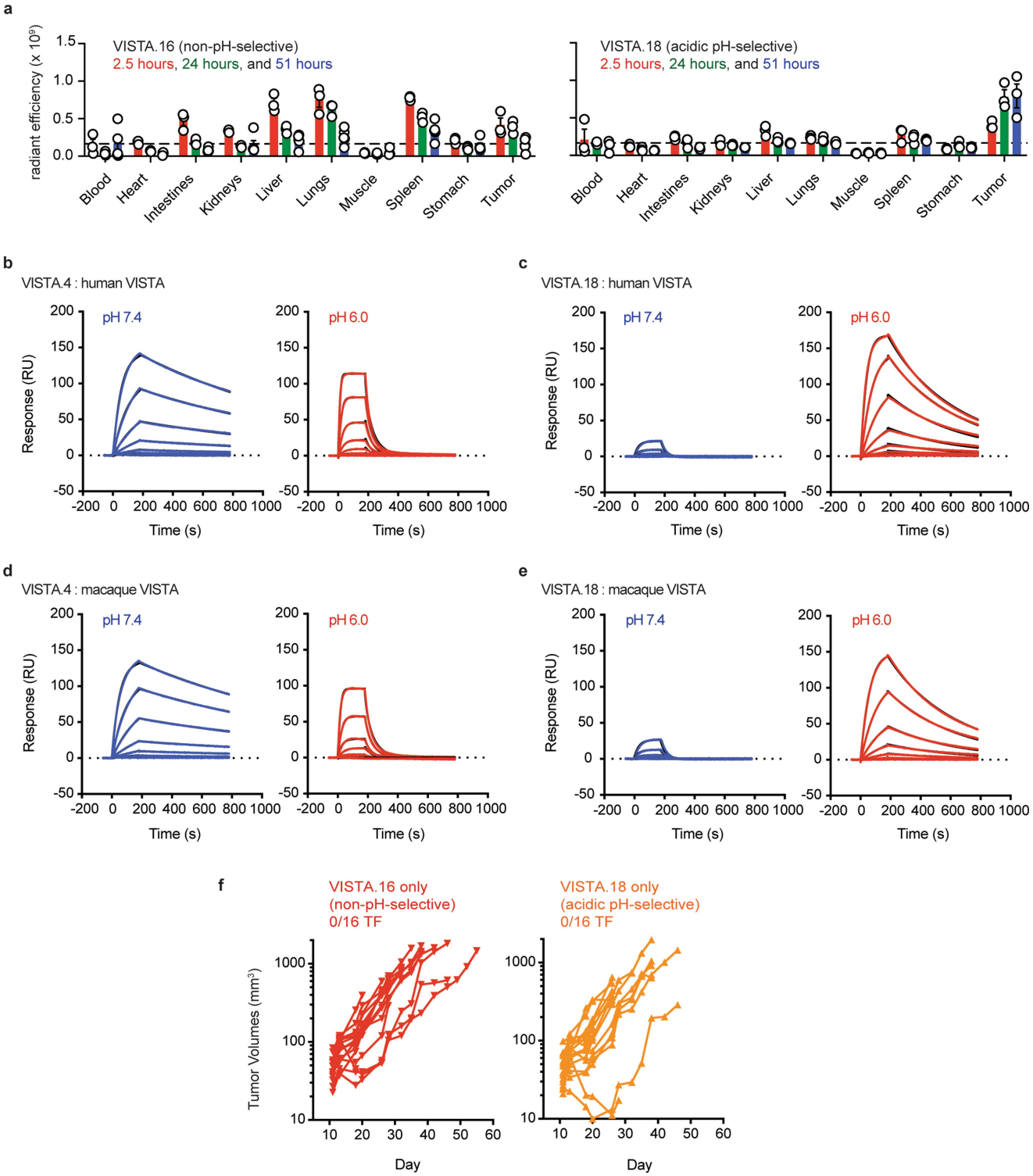



**Extended Data Fig. 8 | Further characterization of the determinants of VISTA–PSGL-1 binding. a**, BLI binding magnitudes of anti-PSGL-1 clone KPL1 to captured total, sulfotyrosine-poor, and sulfotyrosine-rich fractions of PSGL-119-mer–Fc at pH 6.0 (green) and pH 7.4 (blue). These data are representative of one experiment. **b**, BLI binding magnitudes of wild-type PSGL-119-mer–Fc (WT, blue) and tyrosine to alanine mutant PSGL-119-mer–Fc (Y→A, green) to captured VISTA–Fc at the indicated pH. These data are representative of one experiment. **c**, BLI binding magnitudes for VISTA.5 (a non-blocking antibody, left) and VISTA.16 (a blocking antibody, right) binding to captured wild-type (WT, black), 153/154/155 histidine to alanine mutant (H→A, red), 153/154/155 histidine to aspartic acid mutant (H→D, blue) and 153/154/155 histidine to arginine mutant (H→R, green) VISTA–Fc proteins at pH 6.0 and pH 7.4. These data are representative of one experiment. **d**, SPR binding %$R_{max}$ values for VISTA.18 Fab binding to captured wild-type, H→D mutant, H→R mutant and H→A mutant VISTA proteins at pH 6.0 (left, green) and pH 7.4 (right, blue). Binding at 25, 50 and 100 nM are indicated by light, medium and dark coloured bars respectively. These data are representative of one experiment. **e**, Wild-type and mutant VISTA–Fc binding to CHO-PSGL-1 cells at pH 6.0. Data are VISTA–Fc MFI and are representative of two independent experiments. **f**, Wild-type and mutant VISTA–Fc suppression of primary T cell NF-kB phosphorylation at pH 6.8. Data are pNF-kB MFI normalized to control. *n* = five T cell donors; these data are representative of two independent experiments. **g**, Additional human VISTA–Fc recombinant proteins were produced with the histidine residues at positions 98 and 100, or a 98, 100, 153, 154 and 155 (quintuple) replaced by arginine. Wild-type and H→R mutant VISTA–Fc binding to CHO-PSGL-1 cells at pH 7.4. Data are VISTA–Fc MFI and are representative of two independent experiments.




**Extended Data Fig. 9** | See next page for caption.

**Extended Data Fig. 9 | Additional analyses of VISTA mouse models. a**, VISTA.10 antibody blockade of mouse VISTA–Fc binding to activated mouse $CD4^+$ T cells at pH 6.0. Binding with no blocking antibody is included as a control. Data are mouse VISTA–Fc MFI and are representative of two independent experiments. **b**–**d**, MC38 tumour-bearing wild-type mice were treated with PD-1 and VISTA blocking antibodies as described in Fig. 4. $n = 5$ per group; these data are representative of three (**b**) or two (**c**, **d**) independent experiments. **b**, The frequency of intratumoural $CD4^+$ T cells seven days after the start of treatment. $^{***}P = 0.0001$. **c**, LAG-3 and TIM-3 MFI on intratumoral $CD8^+$ T cells. $^{***}P < 0.0001$. **d**, The frequencies of intratumoral leukocytes ($CD45^+$, first plot from the left), macrophages ($CD11b^+MHCII^+LyC^{low}LyG^{low}$), monocytes ($CD11b^+LyC^{high}LyG^{low}$) and granulocytes ($CD11b^+LyC^{low}Ly6G^{high}$). Per cent $CD45^+$, $^{***}P = 0.0001$; per cent macrophages, $^{**}P = 0.0071$; $^{***}P = 0.0009$. **e**, MC38 tumour-bearing VISTA-knockout mice and their wild-type littermates were treated with control or PD-1 blocking antibodies. Frequencies of intratumoral $CD8^+$ (left) and $CD4^+$ (right) T cells seven days after the start of treatment. Per cent $CD8^+$, $^*P = 0.0285$; per cent $CD4^+$, $^*P = 0.0330$. $n = 4$ (KO mice treated with anti-PD-1) or 5 mice per group; these data are representative of two independent experiments. **f**, Schematic of the endogenous (top) and humanized (bottom) VISTA sequence. **g**, Representative Southern blot of 4 heterozygous mice and 1 wild-type mouse (WT) for the humanized *VISTA* allele (7.1 kb) and the endogenous *Vista* allele (6.1 kb). These data are representative of three independent experiments. **h**, Expression of mouse and human VISTA on leukocytes from wild-type (blue) and homozygous human VISTA knock-in mice (red). Non-$T_{reg}$ $CD4^+$ T cell ($CD3^+CD4^+FoxP3^-$), $CD4^+$ $T_{reg}$ cell ($CD3^+CD4^+FoxP3^+$), $CD8^+$ T cell ($CD3^+CD8^+$) and myeloid cell ($CD3^-B220^-CD11b^+$) subsets were assessed. Data are per cent VISTA-expressing and are representative of four independent experiments. **i**, Wild-type mice were treated with a single intravenous injection of 200 μg of an anti-mouse VISTA antibody (red downward triangles) or an isotype-matched control antibody (blue squares). Data are serum antibody concentrations and are representative of two independent experiments. Statistics depict mean ± s.e.m. and one-way ANOVA with Dunnett's multiple comparisons.




**Extended Data Fig. 10 | Additional analyses of human VISTA antibodies in mice and macaques. a**, Quantitative tissue biodistribution of fluorescently labelled VISTA.16 (left) and VISTA.18 (right) at 2.5 h (red), 24 h (green) and 51 h (blue) after injection into MC38 tumour-bearing human VISTA knock-in mice. $n = 5$ (VISTA.16 at 51 h) or 3 mice per group. Data are radiant efficiency mean ± s.e.m. and are representative of two independent experiments. **b**–**e**, Human and cynomolgus macaque sensorgrams for the antibodies VISTA.4 and VISTA.18 at pH 7.4 (blue, left) and pH 6.0 (red, right). These data are representative of two independent experiments. **b**, Human VISTA sensorgrams for VISTA.4. **c**, Human VISTA sensorgrams for VISTA.18. **d**, Cynomolgus macaque VISTA sensorgrams for VISTA.4. **e**, Cynomolgus macaque VISTA sensorgrams for VISTA.18. **f**, MC38 tumour-bearing human VISTA knock-in mice were treated as described in Fig. 4. Tumour growth in mice treated with VISTA.16 only (left) or with VISTA.18 only (right). $n = 16$ mice per group. Data are tumour volumes and are a composite of two independent experiments.

nature research

Corresponding author(s): Robert Johnston

Last updated by author(s): 9/9/2019

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☐ | ☒ | A description of all covariates tested |
| ☐ | ☒ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ | For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☒ | ☐ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒ | ☐ | Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |

*Our web collection on statistics for biologists contains articles on many of the points above.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | Flow cytometry data were acquired with FACSDiva (BD, v 8.0.1) and with CytExpert (Beckman Coulter, v 2.1.0.92).<br>Luciferase data were acquired with EnVision Manager (PerkinElmer, v 1.14.3049.528).<br>Mass spectrometry data were acquired and analyzed with Byonic (Protein Metrics, v3.5).<br>Octet data were acquired and analyzed with ForteBio (Molecular Devices, v 9.0.0.48).<br>SPR data were acquired with Biacore T200 Control Software (GE Healthcare, v 2.0.2).<br>ELISA data were acquired and analyzed by Softmax Pro (Molecular Devices, v 7).<br>Crystal structure data were acquired and analyzed by HKL2000 (HKL Research, v 719).<br>ITC data were acquired with MicroCal Auto-iTC200 (Malvern Panalytical v 1.21).<br>Tissue imaging data were acquired and analyzed with Living Image (PerkinElmer, v 4.5). |
| Data analysis | Flow cytometry data were analyzed by FlowJo (BD, v 10.5.3).<br>SPR data were analyzed with Biacore T200 Evaluation Software (GE Healthcare, v 3.1).<br>ITC data were analyzed by MicroCal PEAQ-ITC (Malvern Panalytical, v 1.21).<br>VISTA : PSGL-1 structural modeling was performed with Maestro (Schrodinger, v 10.7.015).<br>PK data were analyzed with Kinetica (ThermoFisher Scientific, v 5.0).<br>Statistics were visualized and generated with Prism (GraphPad, v 8.0.2). |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Data generated or analyzed in this study are included in the manuscript. The VISTA crystal structure described in Figure 2 has been deposited with wwPDB and will be released upon publication.

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Sample size | No sample size calculations were performed. Sample sizes were modeled after those from past experience and publications. |
| Data exclusions | Data were only excluded from assay development and technically failed experiments. In mouse studies, mice that were not enrolled into a treatment group or who reached an endpoint other than study termination or tumor progression (such as morbidity or tumor ulceration) were excluded. |
| Replication | All technically successful replicate studies reproduced the indicated results. The number of replicates for each study is indicated. Some studies (eg, cynomolgus macaque pharmacokinetics and imaging studies) have only been performed once. |
| Randomization | In mouse studies, mice were randomized at the start of treatment on the basis of tumor volume. |
| Blinding | Investigators were not blinded. Data reported for mouse studies (tumor measurements and flow cytometry) are non-subjective. |

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

| n/a | Involved in the study |
|---|---|
| ☐ | ☒ Antibodies |
| ☐ | ☒ Eukaryotic cell lines |
| ☒ | ☐ Palaeontology |
| ☐ | ☒ Animals and other organisms |
| ☐ | ☒ Human research participants |
| ☒ | ☐ Clinical data |

### Methods

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ ChIP-seq |
| ☐ | ☒ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |

## Antibodies

| | |
|---|---|
| Antibodies used | Commercially sourced antibodies and other flow cytometry reagents:<br>anti-mouse IgG, polyclonal, Alexa Fluor 647, 1:250, Jackson Immunoresearch 115-605-071 lot 131089<br>anti-mouse CD3, 145-2C11, PE-Cy7, 1:200, Biolegend 100320 lot B268542<br>anti-mouse CD4, GK1.5, Brilliant Violet 711, 1:400, Biolegend 100447 lot B245638<br>anti-mouse CD8a, 53-6.7, Brilliant Violet 786, 1:400, Biolegend 100750 lot B273618<br>anti-mouse CD11b, M1/70, PerCP-Cy5.5, 1:200, eBioscience 45-0112-82 lot 1929457<br>anti-mouse CD19, 6D5, APC-Cy7, 1:400, Biolegend 115530 lot B253924<br>anti-mouse CD45, 30-F11, Brilliant Violet 421, 1:500, Biolegend 103133 lot B263588<br>anti-mouse F4/80, BM8, Brilliant Violet 785, 1:200, Biolegend 123141 (lot unavailable)<br>anti-mouse FoxP3, FJK-16s, FITC, 1:100, eBioscience 11-5773-82 lot 2007700<br>anti-mouse Gr1, Rb6-8C5, APC, 1:400, BD Biosciences 553129 lot 7121540 |

anti-mouse LAG-3, eBioC9B7W, PerCP-eFluor 710, 1:200, eBioscience 46-2231-82 lot E15914-105
anti-mouse Ly6C, HK1.4, Brilliant Violet 711, 1:400, Biolegend 128037 lot B247973
anti-mouse MHC-II, M5/114.15.2, APC-Cy7, 1:200, Biolegend 107628 (lot unavailable)
anti-mouse PD-1, J43, APC-eFluor 780, 1:200, eBioscience 47-9985-82 lot 1999029
anti-mouse TIM-3, RMT3-23, PE-Cy7, 1:200, eBioscience 25-5870-82 lot 4342910
anti-mouse VISTA, MIH63, PE, 1:200, Biolegend 150204 (lot unavailable)
anti-human IgG, polyclonal, Alexa Fluor 647, 1:250, Jackson Immunoresearch 709-605-149 lot 129954
anti-human CD3e, SK7, PE, 1:10, Biolegend 344806 lot B246230
anti-human CD4, SK3, Brilliant Violet 650, 1:10, BD Biosciences 563875 lot 7048624
anti-human CD8a, RPA-T8, Brilliant Violet 785, 1:10, Biolegend 301046 lot B221662
anti-human CD14, M5E2, Brilliant Violet 421, 1:10, Biolegend 310830 lot B262218
anti-human CD15, MMA, APC, 1:10, eBioscience 17-0158-42 (lot unavailable)
anti-human CD19, HIB19, APC-R700, 1:10, BD Biosciences 564977 lot 7160751
anti-human CD42b, HIP1, APC, 1:10, Biolegend 303912 (lot unavailable)
anti-human CD56, NCAM 16.2, Brilliant Violet 421, 1:10, BD Biosciences 562752 lot 5295681
anti-human NFkB pS529, K10-895.12.50, unconjugated, 1:100, BD Biosciences 558393 (lot unavailable)
anti-human PSGL-1, KPL1, APC, 1:100, BD Biosciences 562758 lot 7125846
anti-human VISTA, B7H5DS8, APC, 1:10, eBioscience 17-1088-42 (lot unavailable)
anti-human VSIG-3, polyclonal, unconjugated, 1:100, R&D Systems AF4915 (lot unavailable)
CellTrace Violet, ThermoFisher C34557 (lot unavailable)
Live/Dead Fixable Aqua, ThermoFisher L34957 (lot unavailable)
Streptavidin Dextramer, no clone, PE, 32 nM, Immudex DX01-PE (lot unavailable)
human VISTA, no clone, biotinylated, 32-890 nM, Acro Biosystems B75-H82E1 (lot unavailable)
human VISTA-Fc, no clone, unconjugated, 10 ug/mL, R&D Systems 7126-B7-050 (lot unavailable)
mouse VISTA-Fc, no clone, unconjugated, 10 ug/mL, R&D Systems 7005-B7-050 (lot unavailable)

Internally generated antibodies:
anti-human VISTA clones VISTA.4, VISTA.5, VISTA.16, and VISTA.18
anti-mouse VISTA clone VISTA.5

Validation

VISTA.4 validation is depicted in Extended Data Fig 2F-H.
VISTA.5 validation is depicted in Extended Data Fig 2F-H.
VISTA.10 validation is depicted in Extended Data Fig 9A.
VISTA.16 validation is depicted in Extended Data Fig 4D.
VISTA.18 validation is depicted in Figure 2, Extended Data Fig 3D, 4D, 10B, and 10D.

## Eukaryotic cell lines

Policy information about cell lines

| | |
|---|---|
| Cell line source(s) | 293T, Raji, Jurkat, MC38, and CHO cell lines were obtained from ATCC and modified as described in the methods. Expi293 cells were obtained from Gibco. |
| Authentication | No cell line authentication was performed. |
| Mycoplasma contamination | All cell lines were screened for mycoplasma and found to be negative prior to use. MC38 tumor cells were additionally screened for other microbial contaminants and found to be negative prior to use in mouse studies. |
| Commonly misidentified lines (See ICLAC register) | No commonly misidentified cell lines were used. |

## Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

| | |
|---|---|
| Laboratory animals | Mice used in the study were males and females aged 6-12 weeks. Wildtype (C57BL6), VISTA knockout, and human VISTA knock-in strains were used. Cynomolgus macaques were males 6-24 months of age. |
| Wild animals | This study did not involve wild animals. |
| Field-collected samples | This study did not involve field-collected samples. |
| Ethics oversight | Animal studies were conducted in compliance with all relevant ethical regulations. Animal studies performed at Bristol-Myers Squibb were approved by the Bristol-Myers Squibb Institutional Care and Animal Use Committee. Animal studies performed at Five Prime Therapeutics were approved by the Five Prime Therapeutics Institutional Care and Animal Use Committee. Animal studies performed at Murigenics were approved by the Bristol-Myers Squibb Animal Welfare Risk Assessment Team and by the Murigenics Institutional Care and Animal Use Committee. |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Human research participants

Policy information about studies involving human research participants

| | |
|---|---|
| Population characteristics | Human blood donors for experiments were anonymous, and no population information was collected. |
| Recruitment | Donors were anonymously recruited from Bristol-Myers Squibb employees. |
| Ethics oversight | Human blood was obtained from a research blood donation program administered by the Bristol-Myers Squibb Occupational Health and Wellness department. The program was operated in compliance with all relevant ethical regulations, and written informed consent was obtained from all donors. |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Flow Cytometry

### Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

### Methodology

| | |
|---|---|
| Sample preparation | Samples were prepared as listed in the methods. |
| Instrument | Fortessa (BD Biosciences) and Cytoflex (Beckman Coulter) instruments were used to acquire flow cytometry data. |
| Software | All flow cytometry data were analyzed with FlowJo software (BD Biosciences) |
| Cell population abundance | *Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.* |
| Gating strategy | *Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.* |

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Article

# Stabilization of chromatin topology safeguards genome integrity

https://doi.org/10.1038/s41586-019-1659-4

Received: 13 September 2018

Accepted: 10 September 2019

Published online: 23 October 2019

Fena Ochs[1], Gopal Karemore[1,5], Ezequiel Miron[2,6], Jill Brown[3], Hana Sedlackova[1], Maj-Britt Rask[1], Marko Lampe[4], Veronica Buckle[3], Lothar Schermelleh[2]*, Jiri Lukas[1]* & Claudia Lukas[1]

To safeguard genome integrity in response to DNA double-strand breaks (DSBs), mammalian cells mobilize the neighbouring chromatin to shield DNA ends against excessive resection that could undermine repair fidelity and cause damage to healthy chromosomes[1]. This form of genome surveillance is orchestrated by 53BP1, whose accumulation at DSBs triggers sequential recruitment of RIF1 and the shieldin–CST–POLα complex[2]. How this pathway reflects and influences the three-dimensional nuclear architecture is not known. Here we use super-resolution microscopy to show that 53BP1 and RIF1 form an autonomous functional module that stabilizes three-dimensional chromatin topology at sites of DNA breakage. This process is initiated by accumulation of 53BP1 at regions of compact chromatin that colocalize with topologically associating domain (TAD) sequences, followed by recruitment of RIF1 to the boundaries between such domains. The alternating distribution of 53BP1 and RIF1 stabilizes several neighbouring TAD-sized structures at a single DBS site into an ordered, circular arrangement. Depletion of 53BP1 or RIF1 (but not shieldin) disrupts this arrangement and leads to decompaction of DSB-flanking chromatin, reduction in interchromatin space, aberrant spreading of DNA repair proteins, and hyper-resection of DNA ends. Similar topological distortions are triggered by depletion of cohesin, which suggests that the maintenance of chromatin structure after DNA breakage involves basic mechanisms that shape three-dimensional nuclear organization. As topological stabilization of DSB-flanking chromatin is independent of DNA repair, we propose that, besides providing a structural scaffold to protect DNA ends against aberrant processing, 53BP1 and RIF1 safeguard epigenetic integrity at loci that are disrupted by DNA breakage.

To study protection of DNA ends in the context of the 3D nuclear architecture, we set out to visualize chromatin occupancy by 53BP1. Although a typical 53BP1 repair focus appears as a homogenous sphere under conventional microscopy[3,4], 3D structured illumination microscopy (3D-SIM)[5,6] revealed an intrinsically organized compartment consisting of between four and seven 53BP1-labelled sub-domains assembled in an ordered, circular fashion around a central interchromatin space (Fig. 1a). Higher-resolution imaging by stimulated emission depletion (STED) microscopy[7] refined that 53BP1 sub-domains span 60–180 nm with a centre-to-centre distance of approximately 140 nm (Extended Data Fig. 1a–e). We name these sub-domains 53BP1 nanodomains (53BP1-NDs) and their higher-order assembly 53BP1 microdomains (53BP1-MDs; Extended Data Fig. 1f). A similar chromatin arrangement was detected with different SIM instruments, reproduced using independent antibodies against 53BP1, and validated by visualization of endogenous 53BP1 tagged with GFP (Extended Data Fig. 1g–k). The 53BP1 patterns mirrored those of phosphorylated H2AX (γH2AX) and overlapped with core histones (Extended Data Fig. 2a–c), consistent with the finding that DSB sites are organized in chromatin nanodomains[8,9]. A typical 53BP1-MD assembled around one active site of DSB repair, exemplified by a single spot of XRCC4 involved in non-homologous end joining or RPA engaged in homology-directed repair (Fig. 1b, Extended Data Fig. 2d–f). 53BP1-MDs formed in both pre- and post-replicative chromatin (Fig. 1a–c), indicating that they represent a general response to DSBs.

Whereas depletion of shieldin subunits (SHLD2, SHLD3) had no effect on the 3D arrangement of 53BP1-decorated chromatin (Fig. 1c), depletion of RIF1 disrupted 53BP1-MDs into disordered and elongated shapes characterized by misaligned 53BP1-NDs (Fig. 1d). We quantified this topological disruption using a custom-designed quantitative nanoscopy texture (QUANTEX) analysis tool, which revealed a significant increase in the mean breadth and principal axis length of 53BP1-MDs (Fig. 1e, f, Extended Data Fig. 3a–c). The topological disruption was reproduced by silencing *RIF1* with multiple small interfering RNAs (siRNAs), by replacing endogenous 53BP1 with a mutant that cannot promote RIF1 recruitment[10], and in several cancer-derived and

[1]Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. [2]Micron Oxford Advanced Bioimaging Unit, Department of Biochemistry, University of Oxford, Oxford, UK. [3]MRC Molecular Haematology Unit, MRC Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, UK. [4]European Molecular Biology Laboratory, Advanced Light Microscopy Core Facility, Heidelberg, Germany. [5]Present address: Modeling and Predictive Technologies, Novo Nordisk A/S, Måløv, Denmark. [6]Present address: Division of Gene Regulation, The Netherlands Cancer Institute, Amsterdam, The Netherlands. *e-mail: lothar.schermelleh@bioch.ox.ac.uk; jiri.lukas@cpr.ku.dk

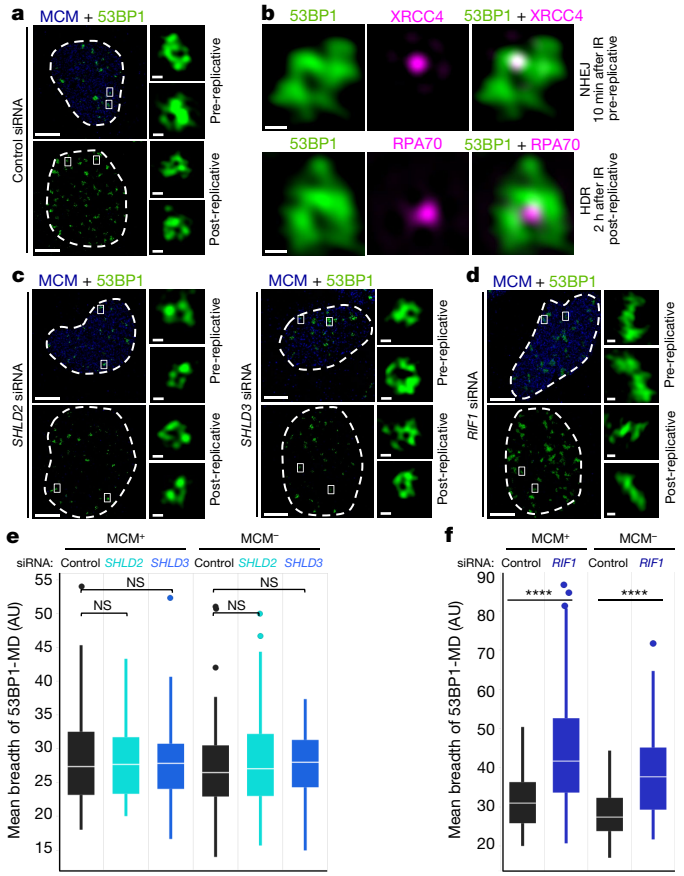



**Fig. 1 | DSBs are surrounded by 53BP1 nanodomains (53BP1-NDs) arranged into higher-order 53BP1 microdomains (53BP1-MDs) in a RIF1-dependent manner. a**, 3D-SIM of GFP–53BP1-MDs in U2OS cells exposed to irradiation (1 Gy, 2 h). MCM, minichromosome maintenance protein. **b**, 3D-SIM of GFP-53BP1-MDs with immunostained XRCC4 (top) or RPA70 (bottom) in U2OS cells exposed to irradiation (IR; 1 Gy) for indicated times. **c**, **d**, 3D-SIM of immunostained 53BP1-MDs after siRNA-mediated depletion of SHLD2 (**c**, left), SHLD3 (**c**, right) and RIF1 (**d**) in cells treated as in **a**. Insets in **a**, **c**, **d** are magnified 53BP1-MDs. **e**, QUANTEX analysis of mean breadth of 53BP1-MDs in cells treated as in **c**; $n = 40$ per condition. **f**, QUANTEX analysis of mean breadth of 53BP1-MDs in cells treated as in **d**; $n = 60$ per condition. Box plots: centre line, median; box limits, 25th and 75th centiles; whiskers, minimum and maximum; dots, outliers. NS, not significant; $P = 0.95, 0.51, 0.60, 0.50$ (**e**, left to right). $^{****}P = 3.8003 \times 10^{-9}$, $1.6698 \times 10^{-9}$ (**f**, left to right). Two-tailed non-parametric Wilcoxon rank-sum test. Cell cycle stage was determined by MCM status ($MCM^+$ pre-replicative; $MCM^-$ post-replicative). Scale bars, 5 µm (**a**, **c**, **d**) and 200 nm (**b** and insets in **a**, **c**, **d**). Experiments were biologically replicated twice with similar results. For detailed image information see Supplementary Table 1.

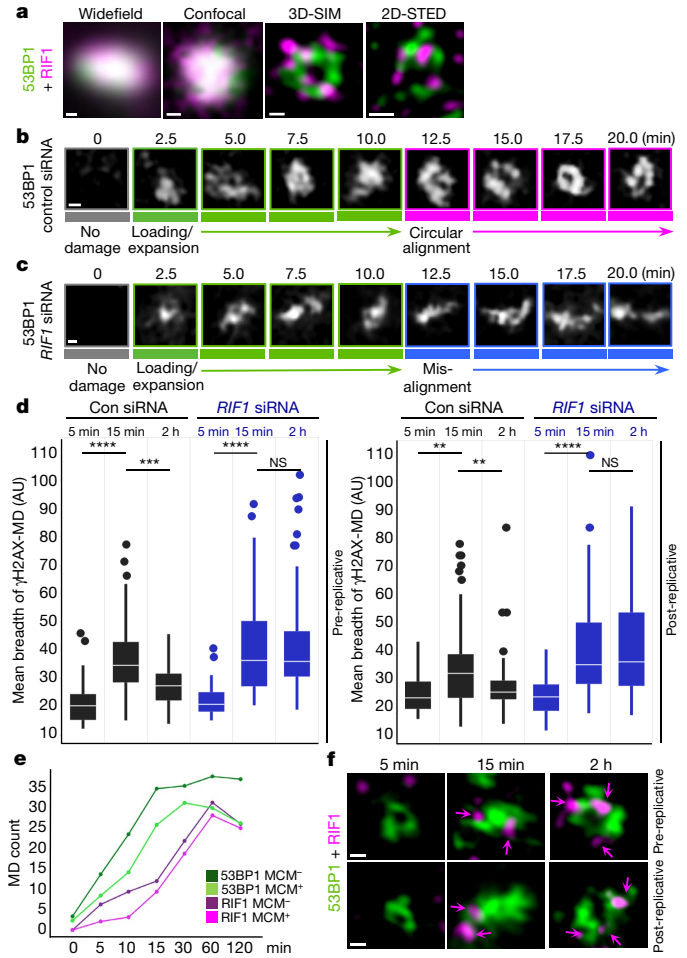



**Fig. 2 | RIF1 localizes to 53BP1-ND neighbourhoods to stabilize ordered and circular architecture of 53BP1-MDs after DNA breakage. a**, GFP–53BP1-MDs in U2OS cells exposed to irradiation (1 Gy, 2 h), immunostained for RIF1 and acquired with conventional (widefield, confocal) or super-resolution (3D-SIM, 2D-STED) microscopy. Pearson correlation coefficient (PCC) is 0.25 ($n = 270$ MDs) showing that colocalization of 53BP1 and RIF1 derived from 3D-SIM was low. **b**, Live-SIM recording of an evolving GFP–53BP1-MD at a single DSB induced by neocarzinostatin (NCS, 10 ng $ml^{-1}$). Manual classification of the main transition is colour-coded. **c**, Live-SIM as in **b** in cells depleted of RIF1. **d**, QUANTEX analysis of mean breadth of γH2AX-MDs in U2OS cells treated with the indicated siRNAs at the indicated times after irradiation (1 Gy); $n = 40$ per condition. Box plots: centre line, median; box limits, 25th and 75th centiles; whiskers, minimum and maximum; dots, outliers. $^{****}P = 8.2676 \times 10^{-12}$, $^{***}P = 1.8363 \times 10^{-04}$, $^{****}P = 1.9056 \times 10^{-08}$, NS (not significant) $P = 0.7366$ (left panel, left to right); $^{*}P = 0.0019$, $^{*}P = 0.0059$, $^{****}P = 3.4337 \times 10^{-09}$, NS $P = 0.9264$ (right panel, left to right); two-tailed non-parametric Wilcoxon rank-sum test. **e**, QIBC analysis of recruitment of 53BP1 and RIF1 to DSBs in cells treated with irradiation (1 Gy) for the indicated times ($n = 500$ cells per condition, data points are means of population). **f**, 3D-SIM of GFP–53BP1-MDs and immunostained RIF1 in U2OS cells treated with irradiation (1 Gy) for the indicated times. Arrows indicate sites of RIF1 recruitment. All scale bars, 200 nm. Experiments were biologically replicated twice (**a**, **d**–**f**) or three times (**b**, **c**) with similar results. For detailed image information see Supplementary Table 1.

non-cancerous cells (Extended Data Fig. 4a–e). Together, these data indicate that 53BP1 and RIF1 form an autonomous module in which RIF1 is required to stabilize 53BP1-NDs into an ordered, circular chromatin architecture (Extended Data Fig. 4f). In support of this concept, knockdown of *TP53BP1* or *RIF1* phenocopied each other by disrupting γH2AX-marked chromatin into disordered and elongated shapes (Extended Data Fig. 4g–i).

To study how 53BP1 and RIF1 cooperate to stabilize chromatin topology, we compared the localization of RIF with that of 53BP1. Whereas conventional microscopy indicates only a general proximity of 53BP1 and RIF1 at DSB sites, 3D-SIM and STED revealed that RIF1 localized to the chromatin boundaries between neighbouring 53BP1-NDs (Fig. 2a). To understand the purpose of this alternating localization, we tracked 53BP1 dynamics from the pre- to post-damaged state using live-cell 3D-SIM (live-SIM; Extended Data Fig. 5a). The first 5–10 min after DSB generation were marked by loading of 53BP1 to DSB-flanking chromatin, aligned with previous findings obtained by conventional microscopy[11]. In the subsequent 5 min, the 53BP1 pattern matured into distinct 53BP1-NDs arranged around a central interchromatin space (Fig. 2b, Extended Data Fig. 5b). The dynamics of 53BP1 was mirrored by γH2AX and Halo-tagged histone H2B (Extended Data Fig. 6a–c), indicating that it was rooted in a chromatin template. Live-SIM analysis of RIF1-depleted

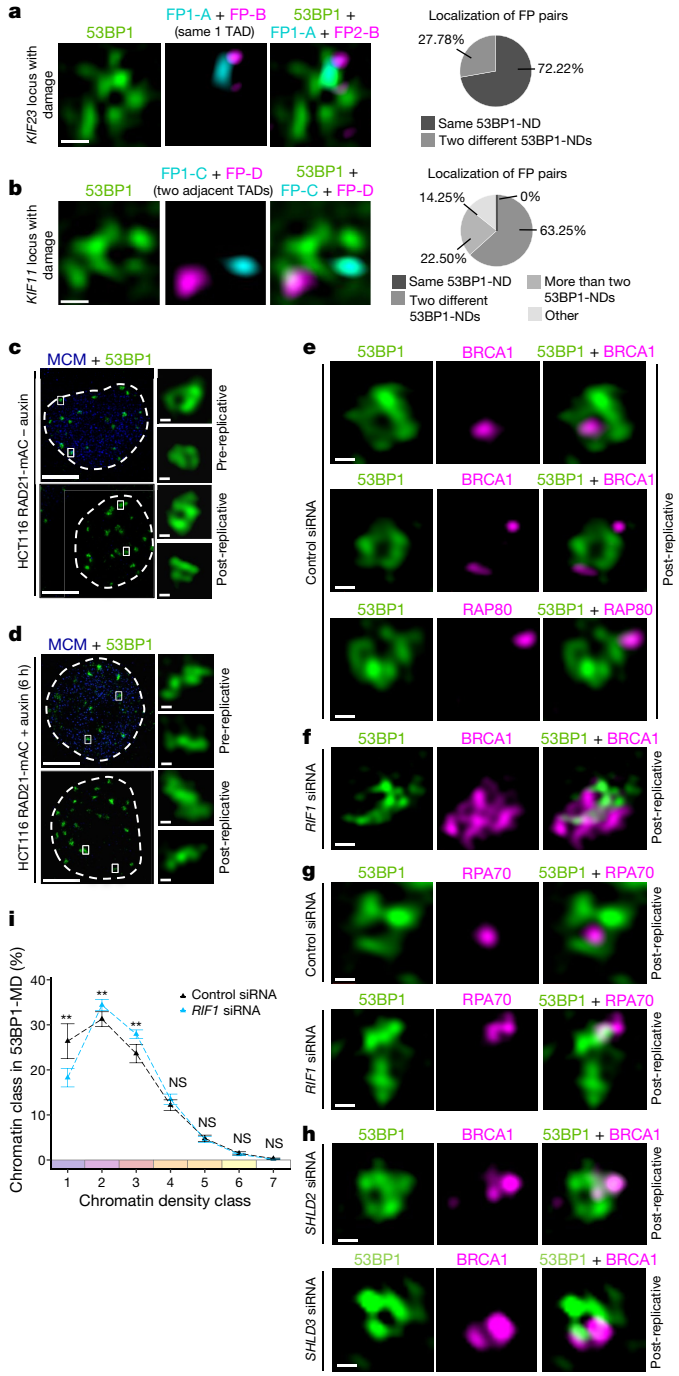



**Fig. 3 | 53BP1-MDs comprise several TAD-sized chromatin domains whose ordered, circular arrangement protects integrity of DSB sites. a**, **b**, 3D-SIM of the KIF23-TAD (**a**; $n = 15$) and the KIF11-TAD (**b**; $n = 41$) labelled with the dual-colour FISH probes (FP-A and FP-B within one TAD; FP-C and FP-C in two TADs); pie charts depict co-localization of the FP pairs with 53BP1-NDs. Other denotes infrequent arrangements. See Extended Data Fig. 8c, d for undamaged TADs. **c**, **d**, 3D-SIM of immunostained 53BP1 in HCT116-RAD21-mAID-mClover cells untreated (**c**) or treated (**d**) with auxin for 6 h. Insets are magnified 53BP1-MDs. **e**, 3D-SIM of GFP–53BP1-MDs in irradiated post-replicative U2OS cells (1 Gy, 2 h), immunostained for BRCA1 or RAP80. Localization frequency within 53BP1-MDs was 28% ($n = 100$) for central BRCA1 (top), 54% ($n = 100$) for peripheral BRCA1 (middle), and 41% ($n = 85$) for peripheral RAP80 (bottom). **f**, 3D-SIM as in top row of **e** after RIF1 depletion. Frequency of aberrantly spread BRCA1 was 85% ($n = 84$). **g**, 3D-SIM of GFP–53BP1-MDs immunostained for RPA70 and treated as in **e**. Localization frequencies were 86% ($n = 92$) for focal RPA70 (top) and 66% ($n = 61$) for elongated RPA70 (bottom). **h**, 3D-SIM as in **e** after depletion of SHLD2 or SHLD3. Frequency of increased but focal BRCA1 was 84% ($n = 119$) for SHLD2 depletion (top) and 73% ($n = 82$) for SHLD3 depletion (bottom). **i**, ChaiN analysis of 53BP1-MDs from wild-type cells or RIF1-depleted cells ($n = 150$ per condition). Medians ± 95% confidence intervals (CI). $^{**}P = 0.0019, 0.0080, 0.0015$ (classes 1–3); NS, $P = 0.1400, 0.6288, 0.2885, 0.1681$ (classes 4–7); two-tailed Student *t*-test. Scale bars, 200 nm (**a**, **b**, **e**–**h** and insets in **c**, **d**); 5 µm (**c**, **d**). Experiments in **c**–**i** were biologically replicated twice with similar results. For detailed image information see Supplementary Table 1.

cells revealed that whereas the initial accumulation of 53BP1 was similar to that in wild-type cells, 53BP1-NDs failed to mature to circular MDs (Fig. 2c, Extended Data Fig. 6d), leading to asphericity of repair foci (quantified as an increase in mean breadth of chromatin marked by γH2AX; Fig. 2d). As quantitative image-based cytometry (QIBC)[12] showed no major change in the levels of γH2AX or chromatin-bound 53BP1 in RIF1-depleted cells, and the number of 53BP1-NDs was not altered when analysed by STED (Extended Data Fig. 7a–d), the likely cause of topological disruptions in RIF1-depleted cells was an inability to stabilize long-range chromatin interactions. Unexpectedly, while these data indicate that 53BP1 and RIF1 cooperate in shaping chromatin architecture around DSBs, QIBC and laser microirradiation[13] independently revealed a temporal difference in their recruitment. In contrast to 53BP1, which was detectable immediately after DNA breakage, RIF1 became discernible only 10–15 min later, when 53BP1-decorated chromatin started to mature into an ordered, circular arrangement (Fig. 2e, f, Extended Data Fig. 7e). Although the shieldin complex resembled RIF1 by localizing to 53BP1-ND neighbourhoods (Extended Data Fig. 7f), disruption of shieldin did not impair the spatial arrangement of 53BP1-NDs (Fig. 1c, e). Thus, recruitment of RIF1 to DSB sites appears to have a unique role in stabilizing chromatin topology initiated by the formation of 53BP1-NDs.

To investigate how the chromatin arrangement at DSB sites influences the general principles of 3D nuclear organization[14], we used CRISPR–Cas9 to introduce single DSBs into TADs that spanned coding sequences for the essential mitotic regulators KIF23 and KIF11 (Extended Data Fig. 8a, b). We then applied RASER-FISH[15] (resolution after single-strand exonuclease resection-fluorescence in situ hybridization), a DNA hybridization technique that complements other TAD-scale approaches[16–19] by allowing the simultaneous detection of labelled FISH probes with super-resolution of immunolabelled proteins. While the labelled TADs showed a similar appearance regardless of DNA damage (Fig. 3a, b, Extended Data Fig. 8c, d), we noticed that the TAD signal in the guide-RNA-targeted loci appeared smaller than the size of the surrounding 53BP1-MDs (Extended Data Fig. 8e, f). Further investigation using 3D-SIM revealed that the labelled *KIF23*-TAD sequence frequently overlapped with a single 53BP1-ND within a given 53BP1-MD (Fig. 3a). When the sequences of two neighbouring *KIF11*-TADs (one targeted by guide RNA and the other free of DNA damage) were labelled, the RASER-FISH signals colocalized with two distinct 53BP1-NDs (Fig. 3b, Extended Data Fig. 8g). Together, these data define a single 53BP1-MD as a 3D multi-TAD assembly. To test whether the observed TAD-like chromatin partitioning might be linked to mechanisms that shape 3D nuclear architecture[20], we knocked down cohesin subunits (RAD21 and SMC1) by siRNA in U2OS cells or depleted RAD21 using an auxin-inducible degron in HCT116 cells. In all conditions, cohesin deficiency phenocopied *RIF1* knockdown by disrupting 53BP1-MDs into disordered, elongated shapes without changing the expression of 53BP1 or γH2AX (Fig. 3c, d, Extended Data Fig. 9a–j). Thus, RIF1 and cohesin cooperate functionally to maintain chromatin topology at sites of DNA breakage.

Disabling non-homologous end joining or homology-directed repair (by inhibiting DNA-PK or depleting CtIP, respectively) did not impair the formation of ordered and circular 53BP1-MDs (Extended Data Fig. 9k–m), raising the possibility that the 53BP1-initiated and RIF1-stabilized

topological arrangement of DSB-flanking chromatin operates as an autonomous 3D structural scaffold for repair reactions. To test this idea, we monitored the localization of BRCA1, a DNA-end processing regulator that counteracts the chromatin-embedded anti-resection barrier[21]. In wild-type settings, BRCA1 was confined to focal compartments either inside or at the periphery of 53BP1-MDs (Fig. 3e). This dual localization is likely to reflect BRCA1 subcomplexes, as only the outer signal, but not the inner signal, could be recapitulated with RAP80 (a component of a BRCA1 sub-complex)[22]. In RIF1-depleted cells, BRCA1 lost its focal appearance due to massive invasion into misshaped chromatin areas (Fig. 3f). This invasion was accompanied by conversion of the highly focal pattern of RPA into elongated structures, indicating excessive DSB resection (Fig. 3g). Whereas depletion of two independent shieldin subunits also increased the localization of BRCA1 to DSB sites, BRCA1 remained confined to foci and the 53BP1-MDs maintained their ordered, circular shape (Fig. 3h). To investigate whether the mislocalization of BRCA1 in RIF1-deficient settings reflects alterations in the underlying chromatin, we quantified histone H2B–GFP occupancy by ChaiN (chain analysis of the in situ nucleome)[23]. Intensity-based segmentation of 3D-SIM images into seven discrete classes of H2B–GFP (Extended Data Fig. 10a), ranging from class 1 (interchromatin space) to class 7 (most compacted heterochromatin), revealed that 53BP1-MDs featured a distinct distribution of chromatin classes. Depletion of RIF1 led this distribution to move towards reduced interchromatin space (class 1) and increased chromatin decompaction (classes 2 and 3; Fig. 3i). As chromatin class distributions in undamaged chromatin remained unchanged in RIF1-depleted cells (Extended Data Fig. 10b), we conclude that RIF1-mediated enforcement of compact chromatin topology is confined to DSB sites.

This study reveals a function for 53BP1 and RIF1 in safeguarding the 3D structure of genomic loci that have been disrupted by DNA breakage (Extended Data Fig. 10c). The ordered topology of DSB-flanking chromatin may function as a barrier to enzymes whose uncontrolled activity could cause collateral DNA and/or chromatin damage. The massive spreading of BRCA1 across the topologically disordered chromatin could be just the one example of structural disruptions that are unleashed in the absence of 53BP1 and RIF1. In addition, the compact structure of 53BP1-MDs might increase the local concentrations of limiting anti-resection factors such as shieldin, which are among the least abundant proteins in the human proteome[24,25] (Extended Data Fig. 10d). Moreover, stabilized chromatin topology could provide a 3D scaffold for physiological DSBs, such as in immunoglobulin diversification. The finding that 53BP1 and RIF1, but not shieldin, are required for long-range chromosomal transactions during immunoglobulin V(D)J recombination[26] is consistent with such a scenario. Finally, as the topological arrangement of DSB-flanking chromatin is independent of DNA repair, and shieldins are phylogenetically younger than the upstream components of the DNA-end protection pathway[24], we speculate that the 53BP1–RIF1 module might have primarily evolved to safeguard epigenetic information encrypted in the 3D chromatin structure that is challenged by DNA breakage.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1659-4.


1. Lukas, J., Lukas, C. & Bartek, J. More than just a focus: The chromatin response to DNA damage and its role in genome integrity maintenance. *Nat. Cell Biol.* **13**, 1161–1169 (2011).
2. Setiaputra, D. & Durocher, D. Shieldin—the protector of DNA ends. *EMBO Rep.* **20**, e47560 (2019).
3. Ochs, F. et al. 53BP1 fosters fidelity of homology-directed DNA repair. *Nat. Struct. Mol. Biol.* **23**, 714–721 (2016).
4. Spies, J. et al. 53BP1 nuclear bodies enforce replication timing at under-replicated DNA to limit heritable DNA damage. *Nat. Cell Biol.* **21**, 487–497 (2019).
5. Demmerle, J. et al. Strategic and practical guidelines for successful structured illumination microscopy. *Nat. Protocols* **12**, 988–1010 (2017).
6. Kraus, F. et al. Quantitative 3D structured illumination microscopy of nuclear structures. *Nat. Protocols* **12**, 1011–1028 (2017).
7. Wegel, E. et al. Imaging cellular structures in super-resolution with SIM, STED and localisation microscopy: a practical comparison. *Sci. Rep.* **6**, 27290 (2016).
8. Chapman, J. R., Sossick, A. J., Boulton, S. J. & Jackson, S. P. BRCA1-associated exclusion of 53BP1 from DNA damage sites underlies temporal control of DNA repair. *J. Cell Sci.* **125**, 3529–3534 (2012).
9. Natale, F. et al. Identification of the elementary structural units of the DNA damage response. *Nat. Commun.* **8**, 15760 (2017).
10. Callen, E. et al. 53BP1 mediates productive and mutagenic DNA repair through distinct phosphoprotein interactions. *Cell* **153**, 1266–1280 (2013).
11. Doil, C. et al. RNF168 binds and amplifies ubiquitin conjugates on damaged chromosomes to allow accumulation of repair proteins. *Cell* **136**, 435–446 (2009).
12. Toledo, L. I. et al. ATR prohibits replication catastrophe by preventing global exhaustion of RPA. *Cell* **155**, 1088–1103 (2013).
13. Bekker-Jensen, S. et al. Spatial organization of the mammalian genome surveillance machinery in response to DNA strand breaks. *J. Cell Biol.* **173**, 195–206 (2006).
14. Marnef, A. & Legube, G. Organizing DNA repair in the nucleus: DSBs hit the road. *Curr. Opin. Cell Biol.* **46**, 1–8 (2017).
15. Brown, J. M. et al. A tissue-specific self-interacting chromatin domain forms independently of enhancer-promoter interactions. *Nat. Commun.* **9**, 3849 (2018).
16. Bintu, B. et al. Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science* **362**, eaau1783 (2018).
17. Cardozo Gizzi, A. M. et al. Microscopy-based chromosome conformation capture enables simultaneous visualization of genome organization and transcription in intact organisms. *Mol. Cell* **74**, 212–222.e215 (2019).
18. Mateo, L. J. et al. Visualizing DNA folding and RNA in embryos at single-cell resolution. *Nature* **568**, 49–54 (2019).
19. Nir, G. et al. Walking along chromosomes with super-resolution imaging, contact maps, and integrative modeling. *PLoS Genet.* **14**, e1007872 (2018).
20. Szabo, Q., Bantignies, F. & Cavalli, G. Principles of genome folding into topologically associating domains. *Sci. Adv.* **5**, eaaw1668 (2019).
21. Densham, R. M. & Morris, J. R. The BRCA1 ubiquitin ligase function sets a new trend for remodelling in DNA repair. *Nucleus* **8**, 116–125 (2017).
22. Sobhian, B. et al. RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites. *Science* **316**, 1198–1202 (2007).
23. Miron, E. et al. Chromatin arranges in filaments of blobs with nanoscale functional zonation. Preprint at https://www.biorxiv.org/content/10.1101/566638v3 (2019).
24. Gupta, R. et al. DNA repair network analysis reveals shieldin as a key regulator of NHEJ and PARP inhibitor sensitivity. *Cell* **173**, 972–988.e923 (2018).
25. Hein, M. Y. et al. A human interactome in three quantitative dimensions organized by stoichiometries and abundances. *Cell* **163**, 712–723 (2015).
26. Ghezraoui, H. et al. 53BP1 cooperation with the REV7-shieldin complex underpins DNA structure-specific NHEJ. *Nature* **560**, 122–127 (2018).

# Methods

### Cell culture

Cells of the human retinal epithelial cell line hTERT-RPE1 (ATCC CRL-4000), BJ fibroblasts (ATCC CRL-2522), HeLa Kyoto cervical cancer cell line (obtained from S. Narumiya), and U2OS osteosarcoma cell line (obtained from Danish Cancer Society) were grown in DMEM containing 10% heat-inactivated FBS and penicillin–streptomycin antibiotics. The following genetically modified cell line were used: U2OS cells stably expressing mouse 53BP1 N-terminally tagged to EGFP (GFP–53BP1)[3] (1 μg/ml puromycin), U2OS cells with endogenous 53BP1 C-terminally tagged with mEGFP (53BP1–GFP), U2OS cells expressing human GFP–53BP1-7A mutant (400 μg/ml geneticin), and U2OS cells expressing GFP–53BP1 and H2B–HaloTag (1 μg/ml puromycin and 400 μg/ml geneticin), U2OS-3×Flag–SHLD3 cells[24] (obtained from C. Choudhary), HeLa H2B–GFP cells (obtained from F. Barr) and human colorectal carcinoma HCT116 cells with an integrated RAD21 degron (RAD21–mAID–mClover)[27] (obtained from E. Lieberman Aiden). HCT116 cells were cultured in McCoy's 5A modified medium with 10% FBS (100 μg/ml hygromycin and 100 μg/ml geneticin). Cells were tested for mycoplasma on a regular basis and authenticated by STR profiling (IdentiCell Molecular Diagnostics).

### Cell lines and plasmids generated for this study

U2OS GFP–53BP1-7A mutant cells were generated using plasmid pAc-GFP-human 53BP1-7A and selection of single clones according to procedures detailed previously[3]. The plasmid was generated by cloning of a PCR fragment from Flag-tagged 53BP1-7A (a gift from A. Shibata) into vector pAc-GFP-C1 and rendered resistant to *TP53BP1* siRNA (Ambion, s14313) using site-directed mutagenesis with the forward primer CTAGAAGACCAGAAAGAGGGTCGCTCAACTAATAAGGAAAATCC. U2OS GFP–53BP1/H2B–HaloTag cells were generated by transfecting the GFP–53BP1 cell line[4] with plasmid H2B–HaloTag and a selection of clones[3]. Plasmid pHTC–histone H2B–HaloTag was generated by cloning a PCR fragment of H2B from an existing H2B–GFP plasmid into the Nhe1 cloning site of pHCT HaloTag CMV-neo vector (Promega, G7711) to generate a C-terminal HaloTag. U2OS cells homozygously expressing C-terminally tagged 53BP1–GFP were generated using CRISPR–Cas9D10A mediated homology-directed repair[28]: cells were transfected with two pX335-U6-Chimeric_BB-CBh-hSpCas9n (D10A) plasmids (Addgene plasmid #42335)[29] expressing Cas9D10A nickase and guide RNAs (antisense: AACACAATCTCCACGATAGC, sense: GTGTAACTGGATTCCTTGCA) and a donor plasmid containing mEGFP flanked by 900-bp homology arms complementary to the C terminus of 53BP1. After 7 days, GFP-positive cells were sorted by fluorescence-activated cell sorting (FACS; Sony SH800Z cell sorter) to obtain a heterozygous population. The homozygously tagged 53BP1–GFP U2OS cell line was obtained by subcloning and validated by western blotting and junction PCR (forward primer: AAGCAGCACCATTCAAGTGC, reverse primer: TCTGGGCCTTCACCTACCTT) followed by Sanger sequencing. The functionality of 53BP1–GFP was tested using DNA damage response readouts.

### Generation of DNA breaks

X-ray irradiation of cells was performed using a XYLON.SMART 160E-1.5 device (160 kV, 6 mA) delivering 11.8 mGy/s. Soft X-rays were filtered using a 3-mm aluminium filter (YXLON International A/S). For laser microirradiation-induced DNA damage[13], cells were seeded on coverslips and treated with 5-bromo-2′-deoxyuridine (24 h, 10 μM, Sigma B9285). The coverslip was mounted on the stage of an inverted Zeiss Axio Observer microscope equipped with a CryLaS pulsed UV-A laser (355 nm), a 40×/0.6 objective and PALM-Robo software (Version 4.5.09, Carl Zeiss MicroImaging). Laser energy output was determined by biological calibration. For temporal analysis, ten fields were irradiated for 2.5 min each along a straight-line pattern and after completion (25 min) the coverslip was immediately fixed in 4% formaldehyde. To generate site-specific DNA breaks, cells were transfected with gRNA–Cas9 ribonucleoprotein complexes using Lipofectamine CRISPRMAX Cas9 (Invitrogen, CMAX00008). CRISPR RNA and trans-activating CRISPR RNA were annealed according to the manufacturer's instructions (Integrated DNA Technologies). For transfection of a 35-mm dish (2 ml), 6.25 μl Cas9 enzyme (TrueCut Cas9 V2, Invitrogen, A36496, 1 mg/mL) was diluted in 100 μl Opti-MEM medium followed by addition of 12.5 μl duplexed gRNA (2 μM) and 12.5 μl Plus-Reagent from the CRISPRMAX kit. CRISPRMAX reagent (7.5 μl) was diluted in 100 μl Opti-MEM medium in a separate tube, mixed with the other components, incubated at room temperature for 15 min and added to cells. To induce DNA DSBs for live-SIM imaging, cells were treated with the radiomimetic neocarzinostatin (NCS) at a final concentration of 10 ng $ml^{-1}$.

### Gene silencing by siRNA

Transfection of siRNAs (Ambion Silencer Select) was performed with Lipofectamine RNAiMAX (Thermo Fisher Scientific, 13778075) at a concentration of 20 nM. siRNAs used targeted *TP53BP1* (#1 s14314, #2 s14313), *RAD21* (s11726), *RIF1* (#1 s30377, #2 s30378), *SMC1A* (#1 s15753, #2 s15751) and *XRCC4* (s14951). Unless stated otherwise, siRNAs #1 were used. siRNA against *RBBP8*[30] has been previously published. Ambion negative control #1 was used as control siRNA.

### Other treatment of cells

The DNA-PK inhibitor NU7441 (Selleckchem) was used at 10 μM, 1 h before irradiation. In order to induce RAD21 degradation in the RAD21–mAID–mClover cell line[27], cells were treated with 500 μM of the auxin component 3-indoleacetic acid (IAA; Sigma, I2886).

### Antibodies for immunofluorescence detection and western blotting

The following antibodies were used: 53BP1 (mouse, Millipore, MAB3802, 1:750 for immunofluorescence (IF)), 53BP1 (rabbit, Novus Biologicals, NB100-305, 1:750 for IF, 1:1,000 for western blotting (WB)), 53BP1 (rabbit, Novus Biologicals, NB100-304, 1:1,000 for WB), BRCA1 (mouse, Calbiochem, O92, 1:100 for IF), CtIP (mouse, Active Motif, 61141, 1:250 for WB), Flag-Tag (mouse, Sigma, F1804, 1:300 for IF), GFP (rabbit, Torrey Pines Biolabs, TP401, 1:1,000 for WB), H2AX phospho-S139 (mouse, Abcam, ab22551, 1:1,000 for IF), H2AX phospho-S139 (rabbit, Cell Signaling, 9733, 1:1,000 for IF), HaloTag (mouse, Promega, G921A, 1:1,000 for WB), H2B (rabbit, Abcam, ab1790, 1:2,000 for WB), KAP1 (rabbit, Bethyl Laboratories, A300-274A, 1:2,000 for WB), MCM2 (mouse, Novus Biologicals, H00004171-M01, 1:200 for IF, 1:1,000 for WB), MCM5 (rabbit, Abcam, ab17967, 1:200 for IF), MCM7 (mouse, Santa Cruz, sc-9966, 1:1,000 for WB), MCMBP (rabbit, Novus Biologicals, NBP1-90746, 1:1,000 for WB), NUDC (rabbit, Sigma-Aldrich, HPA027183, 1:1,000 for WB), RAD21 (mouse, Millipore, 05-908, 1:500 for WB), RAP80 (Bethyl Laboratories, A300-764A, 1:400 for IF), RIF1 (rabbit, Bethyl Laboratories, A300-569A, 1:500 for IF), RIF1 (rabbit, Cell Signaling, 95558, 1:500 for IF, 1:1,000 for WB), RPA70 (rabbit, Abcam, ab79398, 1:300 for IF), SMC1 (rabbit, Novus Biologicals, NBP2-67733, 1:1,000 for WB), tubulin (mouse, Santa Cruz, SC-8035, 1:500 for WB), XRCC4 (rabbit, Abcam, ab213729, 1:100 for IF). MCM2 (mouse monoclonal) and MCM5 (rabbit polyclonal) antibodies were used to identify pre- and post-replicative cells. Secondary-antibody conjugates for immunofluorescence staining (IF) were goat anti-mouse and goat anti-rabbit Alexa Fluor 488 (A11029, A11034), Alexa Fluor 568 (A11031, A11036) and Alexa Fluor 647 (A21236, A21245) reagents (Invitrogen, highly cross-adsorbed). Secondary-antibody conjugates for STED were goat anti-mouse and anti-rabbit STAR RED (Abberior, 2-0002-011-2, 2-0012-011-9) and goat anti-mouse and anti-rabbit STAR 580 (Abberior, 2-0002-005-1, 2-0012-005-8). For imaging of fixed HeLa cells for H2B–GFP by 3D-SIM, a GFP booster was used (Chromotek, gba488, 1:200). For live-SIM, H2B–HaloTag-expressing cells were labelled with 200 nM Janelia Fluor 585 HaloTag ligand (gift from L. Lavis) 20 min before image acquisition.

### Western blotting

Detection of proteins by western blotting was done using standard procedures and ECL-based chemiluminescence detection. For gel source data, see Supplementary Fig. 1.

### Immunofluorescence staining

Procedure for standard IF has been previously described[3]. IF for 3D-SIM was adapted from previously published protocols[5,31]. In brief, cells were grown on square 18 × 18-mm or 22 × 22-mm #1.5H high-precision coverslips (Marienfeld Superior, thickness 0.170 ± 0.005 mm), rinsed in PBS, pre-extracted, or not, in ice-cold 0.2% PBS-Triton-X for 1 min on ice, as indicated in Supplementary Table 1, and fixed in 4% formaldehyde for 15 min. Primary and secondary antibodies were diluted in antibody diluent (DMEM medium containing 10% FBS and 0.05% sodium azide, filtered through a 0.2-µm filter). Coverslips were washed in distilled water and mounted on a 30-µl drop of non-hardening Vectashield (Vectorlabs, H-1000) or non-hardening Slowfade Diamond (Thermo Fisher Scientific, S36963). For DAPI staining, secondary antibody solution was supplemented with 4′,6′-diamidino-2-phenylindole-dyhydrochloride (DAPI, 0.5 µg/ml).

### FISH probes and labelling

FISH probes (FP) were generated by labelling bacterial artificial chromosomes (BACs, BACPAC Resources Center, https://bacpacresources.org/) with fluorescent dyes. To detect the TAD that harbours the *KIF23* gene as annotated in the ensemble-annotated Hi-C resource at 10-kb resolution (3D Genome Browser, YUE Laboratory, http://promoter.bx.psu.edu/hi-c/view.php), we used two adjacent FISH-BAC probes. KIF23 FP-A is RP11-347N18, labelled with Alexa Fluor 647-aha-dUTP (A32763, Invitrogen); KIF 23 FP-B is RP11-1150H19, labelled with Alexa Fluor 594 5-dUTP (C11400, Invitrogen), together spanning nearly the entire TAD (hg19:chr15: about 69300000–69750000). The FISH-BAC probe FP-C for detecting the TAD that harbours the *KIF11* gene (hg19:chr10: about 94250000–94650000) was BAC probe RP742C13, labelled with Alexa Fluor 647-aha-dUTP. The FISH-BAC probe FP-D for the adjacent TAD (hg19:chr10: about 94650000–95050000), was RP81C11, labelled with Alexa Fluor 594 5-dUTP. Comparison of these TADs in other cell lines and other data sets using the Compare Hi-C function of the YUE laboratory website (http://promoter.bx.psu.edu/hi-c/view.php) showed that they align across different cell lines and Hi-C resolution scales. BAC probes were directly labelled by nick translation as described previously[15].

### RASER-FISH

RASER-FISH maintains nuclear fine-scale structure by replacing heat denaturation with exonuclease III digestion of one of the two DNA strands after UV-generation of nicks and is suitable for super-resolution image analysis. RASER-FISH was conducted as previously described[15] and here was combined with site-specific DSB generation and IF staining of 53BP1 to allow visualization of TADs at sites of damage. As a counterpart to TADs with DSBs, undamaged TADs (Extended Data Fig. 8c, d) were selected by absence of a 53BP1 signal in the volume. In brief, U2OS cells were seeded on 22 × 22 mm #1.5H high-precision coverslips (thickness 0.170 ± 0.005) and labelled for 24 h with 10 µM BrdU/BrdC mix (3:1). Site-specific DSBs were induced by transfection with gRNAs for *KIF23* or *KIF11* (Integrated DNA Technologies, Hs.Cas9.KIF23.1.AB; Hs.Cas9.KIF11.1.AA) as described above. Three hours after gRNA transfection, cells were fixed with 4% formaldehyde (prepared from 16% formaldehyde EM grade ampules) and stained for 53BP1 as described above. After incubation with DAPI for UV sensitization (0.5 µg/ml, 15 min), cells were treated with UV light (254 nm, 15 min) and exonuclease III (NEB, 5 U/µl at 37 °C, 15 min). Labelled probes were denatured in hybridization mix (90 °C, 10 min) and pre-annealed with human Cot-1 DNA (Invitrogen, 37 °C, 15 min) and used for hybridization (39 °C, overnight). Coverslips were washed twice in 1× SSC (37 °C, 30 min) and once in 1× SSC at room temperature. Coverslips were washed in PBS, post-fixed in 4% formaldehyde for 10 min, rinsed in PBS and MilliQ water and mounted in Slowfade Diamond.

### Microscopy and image analysis

Detailed information on all images (imaging modalities, microscopy setups, fluorophores, image processing, display and analysis) can be found in Supplementary Table 1. Image acquisition for QIBC by high-content Widefield microscopy (ScanR Screening station, Olympus) was performed as previously described[4,12]. Images were processed and analysed using ScanR analysis software (Olympus, 2.6.1). Metrics for the different objects (number and intensities of nuclei and foci) were quantified with single and calculated parameters. These values were then exported and visualized using TIBCO Spotfire desktop software (version 7.8.0). To visualize overlapping markers, low *y*-axis jittering was applied in scatter plots (random displacement of objects along *y*-axis). Confocal imaging was carried out on a LSM 880 microscope (Zeiss) or a UltraView Vox spinning disk system (Perkin Elmer). Super-resolution 3D-SIM imaging was carried out following previously described protocols[5], using an ELYRA PS.1 microscope system (Zeiss) and a DeltaVision OMX V3 Blaze system (GE Healthcare). Computational image reconstruction for ELYRA PS.1 was done using theoretical optical transfer functions (OTFs) and the Zeiss algorithm (ZEN BLACK). For OMX V3 Blaze, raw data were reconstructed using channel-specific OTFs[5] (SoftWoRx 6.1). See Supplementary Table 1 for detailed description of imaging modalities, image processing and quality controls by SIMcheck[32]. Live-SIM was carried out on the DeltaVision OMX V3 Blaze system. Cells were seeded in 35-mm glass-bottom dishes (thickness 170 µm ± 5 µm; Ibidi), labelled with 200 nM Janelia Fluor 585 HaloTag ligand (gift from L. Lavis) 20 min before image acquisition and washed in imaging medium (DMEM, Gibco 31053028). To induce DNA DSBs, cells were treated with NCS (10 ng ml$^{-1}$). Samples were imaged at 37 °C and 5% $CO_2$ using an Olympus 60×/1.42 NA PlanApo N objective and RI 1.520 immersion oil. 3D-SIM stacks were acquired over a 0.875-µm-thick (7 *z*-planes) nuclear mid-section to minimize bleaching. To increase throughput, 5–10 nuclei were marked per run and 15 raw images per plane were acquired per time-point and position. The raw data were computationally reconstructed using SoftWoRx 6.1 (GE Healthcare) using channel-specific OTFs as specified in Supplementary Table 1. For analysis and display, only those examples were selected that could be tracked from before to after damage, stayed in focus and did not bleach more than 30% during the whole acquisition. STED imaging was performed on an Abberior STED and RESOLFT 775 QUAD scanning microscope (Abberior Instruments GmbH) using the 488 nm CW laser and 594 nm, and 640 nm pulsed excitation lasers, and a pulsed 775 nm STED laser for depletion using a 100×/1.4 NA oil immersion objective and a 2D depletion donut for enhancing lateral resolution to approximately 50 nm. STED data were analysed and quantified using Fiji/Image J[33].

### 3D image analysis using in-house-developed QUANTEX software

QUAntitative Nanoscopy TEXture analysis (QUANTEX) is a custom image analysis software tool with a graphical user interface, developed in Matlab (R2018a, Mathworks Inc.) to analyse complex 3D cellular structures. The QUANTEX software, manual and webinar can be downloaded from https://figshare.com/s/46fa39d1010d77f51d9c. QUANTEX uses 3D slice-by-slice segmentation followed by connecting segmented components in 3D. Objects are segmented by processing and segmentation algorithms, morphology filtering and advanced watershed algorithms and analysed by original (in-house) and MathWorks algorithms for texture, geometry and morphology features. For segmentation of nuclei, *z*-stacks were clipped to the minimum number of slices, smoothened by Gaussian filter blurring, and then underwent automated weighted Otsu-based segmentation. 53BP1-MDs were segmented in this order: nuclear background subtraction (Rolling Ball size 3), automated Otsu segmentation, morphology filtering (minimum object size 10 voxels).

The parameter output of primary and secondary object features was exported as .xlsx document. The two main QUANTEX features used in this study are principal axis length and mean breadth. The principal axis length feature was implemented in QUANTEX from MathWorks (R2018a, MathWorks Inc.) and is a standard metric for the length of the major axis of an ellipsoid. Mean breadth is a metric from integral geometry and was implemented to QUANTEX from

https://github.com/mattools/matImage/blob/master/matImage/imMinkowski/imMeanBreadth.m. The algorithm computes the integral of mean curvature as a Minkowski measure that is estimated from the Crofton formula (see detailed information in the QUANTEX manual and webinar; https://figshare.com/s/46fa39d1010d77f51d9c). Steps for calculating mean breadth from 3D binary object are as follows: i) Calculate the number of voxels within the object (nv); ii) Calculate the number of connected components in three main direction *x*, *y*, and *z* (ncx, ncy, ncz); iii) Calculate the number of square faces on the plane with normal direction *x*, *y* and *z* (nfx, nfy, and nfz); iv) Calculate mean breadth (MB) in *x* direction MBx = nv − (ncy + ncz) + nfx, *y* direction MBy = nv − (ncx + ncz) + nfy, *z* direction MBz = nv − (ncx + ncy) + nfz; Mean breadth of an object = (MBx + MBy + MBz)/3. Principal axis length and mean breadth each measure the maximum linear dimension of 3D objects. Both measures consistently give significant *P* values and robustly discriminate between globular and elongated 53BP1-MDs. A Spearman's correlation score (test of association between both measures) of $R^2 = 0.59$ (Extended Data Fig. 3c) shows that they carry similar but not identical information: 59% of variation in mean breadth is explained by principal axis length and 41% of variation in mean breadth is independent of the latter. Wilcoxon tests show that mean breadth more robustly discriminates between globular and elongated shapes of 53BP1-MDs and is less susceptible to geometrical outliers; for these reasons, it was chosen as the main measure in this study.

### Image analysis for ChaiN method

This image analysis pipeline was used to extract chromatin density distribution within 53BP1-MDs in an automated manner[23]. Reconstructed and aligned multichannel 3D-SIM micrographs of chromatin and 53BP1-MDs were split into their single channel components and 53BP1-MDs were thresholded by Otsu algorithm and by size exclusion (excluding signal from antibody noise). The H2B chromatin channel was segmented into seven arbitrary classes implementing a hidden Markov model (HMM), where class 1 denotes no detectable chromatin (interchromatin compartment, IC), and classes 2–7 denote increasing levels of chromatin compaction[34]. The 53BP1-MD volumes are used to mask the segmented chromatin, giving the distribution of chromatin density within these volumes. Aggregating these distributions over all sub-volumes for all images yields an average distribution for each density class as a percentage within class-specific statistical confidence ranges. As a control, the whole nuclear volume can also be taken to analyse whether the chromatin distribution changes genome-wide, outside 53BP1-MDs. This workflow runs on free and open source software (Octave and R). Scripts used can be found at https://github.com/ezemiron/Chain.

### RNA sequencing data source

RNA sequencing (RNA-seq) data for *TP53BP1*, *RIF1* and *SHLD1* transcripts were derived from publicly available RNA-seq data sets at EMBL-EBI expression atlas (https://www.ebi.ac.uk/gxa/home). Original data sources are: NIH Genomic Data Commons Cell lines CCLE osteosarcoma (U2OS), Sanger Genomics of Drug Sensitivity in Cancer Project GDSC Cancer Genome Project uterine cervix/cervical carcinoma (HeLa #1), 675 Genentech uterine cervix/cervical adenocarcinoma (HeLa #2), RNA-seq of long poly-adenylated RNA and long non-polyadenylated RNA from ENCODE cell lines/total RNA/whole cell (IMR90) and Genentech RNA seq of 675 commonly used human cancer cell lines (HBL100, breast, normal at time of derivation).

### Statistics and reproducibility

The two-tailed Student's *t*-test was used to test Gaussian distributed per-class data in ChaiN analysis. The two-tailed non-parametric Wilcoxon rank-sum test for equal medians was used for all data underlying box plots except in Extended Data Fig. 7d. Here, the Cochran–Armitage chi-square test was applied to compare the frequency distribution of an ordinal variable between different conditions. Spearman's correlation coefficients and their $R^2$ values were calculated for metrics mean breadth and principal axis length derived from control (negative class) and RIF1 depletion data (positive class) and combined in order to test the association between the metrics. The Pearson correlation coefficient was used to quantify the degree of colocalization between two fluorophores. Experiments were not randomized and no blinding was used during data analysis. Sample size was not pre-determined. Sample size, statistical tests and the number of biological replicates for each experiment are indicated in the figure legends.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Code availability

Custom ChaiN code is available at https://github.com/ezemiron/Chain. Custom QUANTEX code is available from the corresponding author upon reasonable request.

## Data availability

Numerical and statistical source data for Figs. 1e, f, 2d, e, 3a, b, e–i and Extended Data Figs. 1d, e, 2c, e, f, 3b, 4d, 5b, 6c, d, 7a, b, d, 8c, d, 9c, 10b, d are provided online. Primary imaging data underlying widefield, confocal, SIM and STED images in Figs. 1a–d, 2a–c, f, 3a–h and Extended Data Figs. 1c, i–k, 2a–c, 4b, c, e, f, h, i, 5b, 6a, c, d, 7c, e, f, 8b–g, 9b, f–h, k, l have been deposited at the European Bioinformatics Institute (EBI) BioStudies database (https://www.ebi.ac.uk/biostudies/) with accession number S-BSST275. Processed imaging data sets underlying QIBC, QUANTEX, ChaiN and other analysis, including guidance on how to navigate data sets, are available from the corresponding authors. There are no restrictions on data availability.


27. Rao, S. S. P. et al. Cohesin loss eliminates all loop domains. *Cell* **171**, 305–320.e324 (2017).
28. Koch, B. et al. Generation and validation of homozygous fluorescent knock-in cells using CRISPR-Cas9 genome editing. *Nat. Protocols* **13**, 1465–1487 (2018).
29. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
30. Sartori, A. A. et al. Human CtIP promotes DNA end resection. *Nature* **450**, 509–514 (2007).
31. Miron, E., Innocent, C., Heyde, S. & Schermelleh, L. In vivo and in situ replication labeling methods for super-resolution structured illumination microscopy of chromosome territories and chromatin domains. *Methods Mol. Biol.* **1431**, 127–140 (2016).
32. Ball, G. et al. SIMcheck: a toolbox for successful super-resolution structured illumination microscopy. *Sci. Rep.* **5**, 15915 (2015).
33. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
34. Smeets, D. et al. Three-dimensional super-resolution microscopy of the inactive X chromosome territory reveals a collapse of its active nuclear compartment harboring distinct Xist RNA foci. *Epigenetics Chromatin* **7**, 8 (2014).



**Acknowledgements** The HeLa Kyoto cell line was a gift from S. Narumiya. The U2OS cell line stably expressing RINN1/SHLD3 was a gift from C. Choudhary. The HCT116 cell line with endogenously integrated RAD21–mAID–mClover was a gift from E. Lieberman Aiden. The HeLa cell line stably expressing Histone H2B–GFP was a gift from F. Barr. The pX335-U6-Chimeric_BB-CBh-hSpCas9n(D10A) plasmid was a gift from F. Zhang. The Flag-human53BP1-7A plasmid was a gift from A. Shibata. The JF585 dye was a gift from L. Lavis. Research funding for the Lukas laboratory was provided by the Novo Nordisk Foundation (grants NNF14CC0001 and NNF16CC0020906). We further acknowledge support from Wellcome Trust Strategic Awards 091911 and 107457 supporting advanced microscopy at the Micron Oxford Advanced Bioimaging Unit (L.S.), Medical Research Council award MC_UU_12009/1 (J.B. and V.B.), and the Advanced Light Microscopy Facility (ALMF) at EMBL, Heidelberg, and Abberior Instruments, Göttingen (M.L). We thank J. Dreier from the Protein

Imaging Platform at Novo Nordisk Center for Protein Research for help with image analysis, and P. H. Varas from the Core Facility for Integrated Microscopy for advice on super-resolution microscopy, K. Somyajit for conceptual input to this study and all members of the Lukas laboratory for suggestions.

**Author contributions** F.O., C.L. and J.L. conceived the project. F.O. carried out all 3D-SIM, live-SIM, STED, and QIBC experiments and corresponding data analysis. G.K. performed statistical tests and developed QUANTEX together with F.O. and C.L. H.S. performed endogenous tagging of 53BP1. L.S. developed live-SIM and supported F.O. with SIM data acquisition, data analysis and interpretation. E.M. provided ChaiN expertise and analysed ChaiN SIM data acquired by F.O. J.B. and V.B. provided RASER-FISH expertise and J.B. supported F.O. with sample preparation, data acquisition and analysis for RASER-FISH. M.L. taught F.O. STED imaging and supported STED data acquisition and interpretation. C.L. designed the site-specific DSB generation. M.-B.R. performed Western blots and generated cell lines. C.L., J.L. and L.S. supervised the project and together with F.O. wrote the manuscript. All authors contributed to manuscript editing.

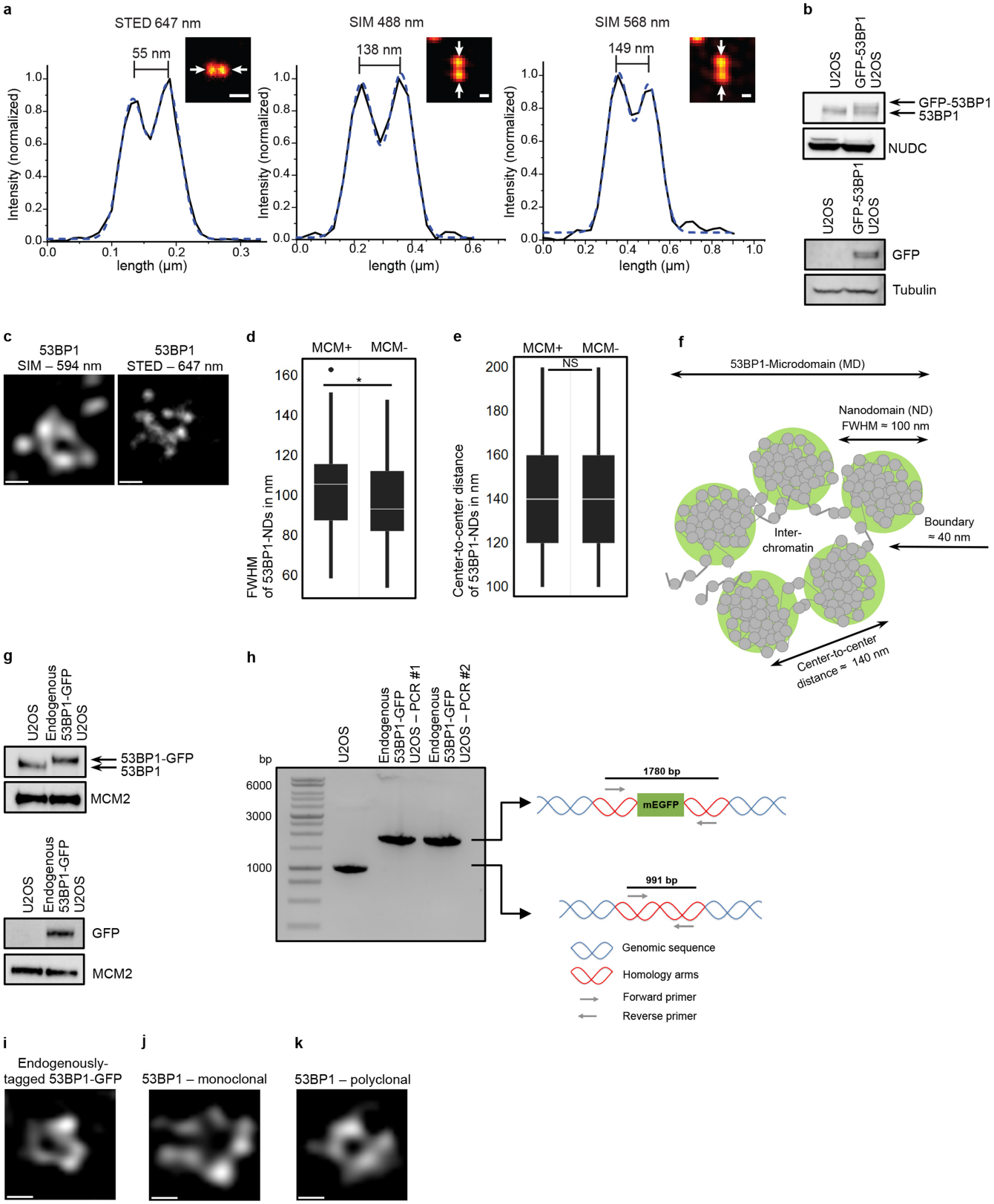
**Competing interests** The authors declare no competing interests.

**Additional information**

**Supplementary information** is available for this paper at https://doi.org/10.1038/s41586-019-1659-4.

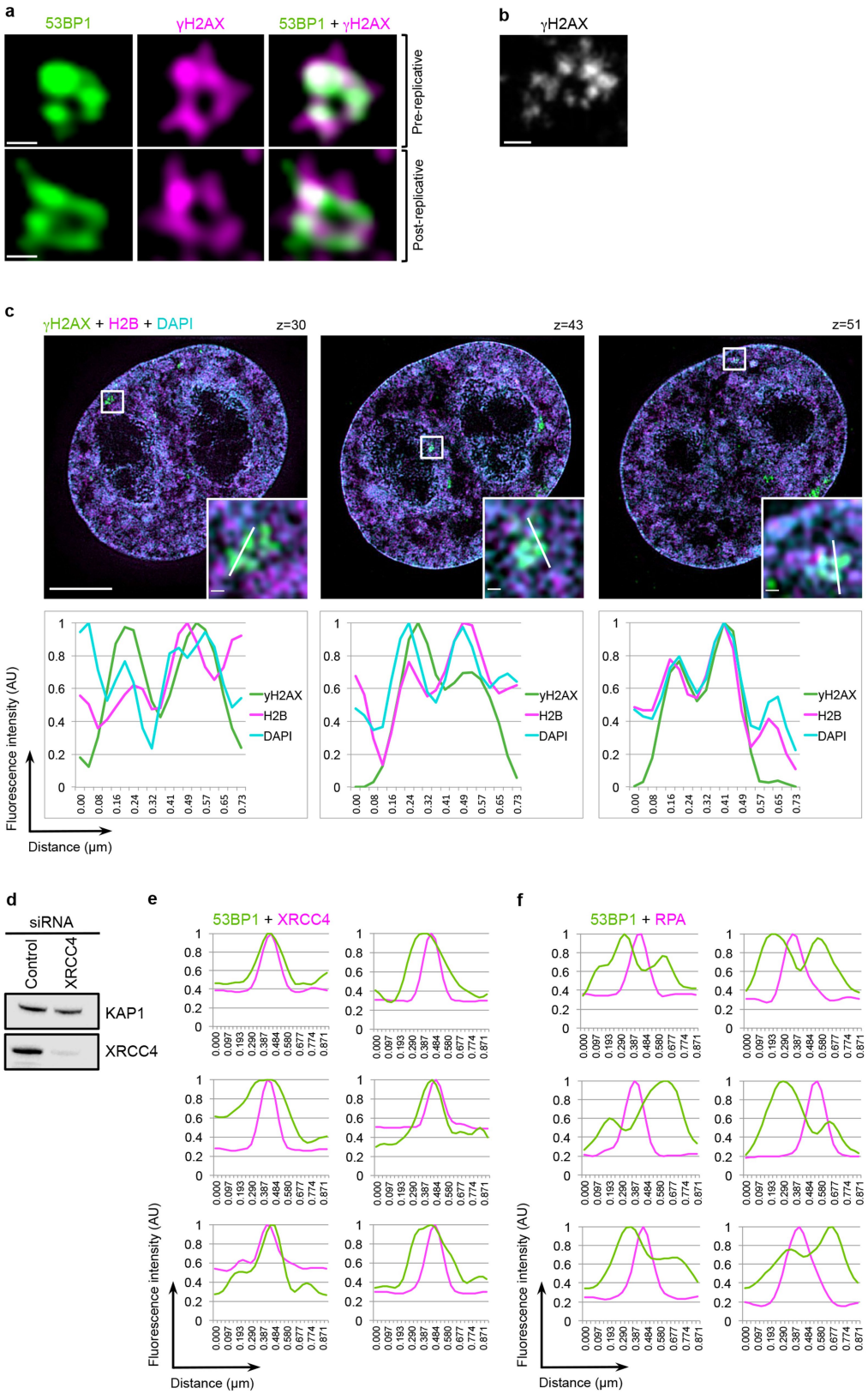
**Correspondence and requests for materials** should be addressed to L.S. or J.L.

**Peer review information** *Nature* thanks Yujie Sun and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.




**Extended Data Fig. 1** | See next page for caption.

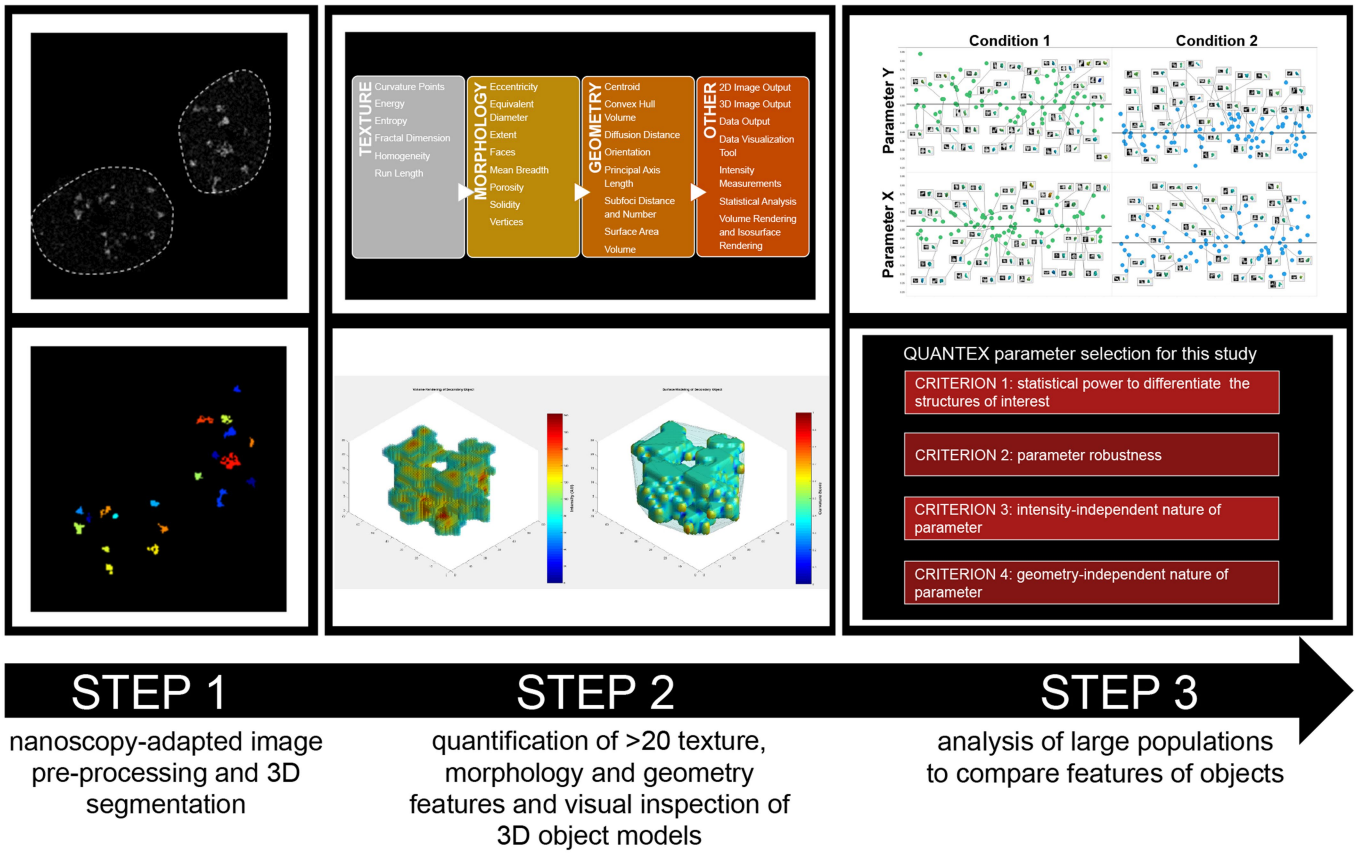
**Extended Data Fig. 1 | Spatial features of 53BP1-MDs at sites of DNA breakage. a**, Experimentally derived resolution for STED and 3D-SIM instruments using nano-bead imaging under identical conditions as for image data acquisition at the indicated excitation wavelengths. Line profile is average of three lines; dotted line shows fit of a double Gaussian distribution, where the peak-to-peak distance indicates spatial resolution. **b**, Western blot of GFP–53BP1 U2OS cells immunostained for 53BP1, GFP and loading controls (NUDC, tubulin). **c**, 3D-SIM and STED images of immunostained 53BP1-MDs in U2OS cells exposed to irradiation (1 Gy, 2 h). Images were processed identically for pixel numbers and bicubic interpolation smoothing for direct comparison. **d**, Diameter of a 53BP1-ND in pre- and post-replicative cells determined by full width at half maximum (FWHM, $n = 75$) from STED data in **c**. **e**, Centre-to-centre peak distance ($n = 85$) of 53BP1-NDs from STED data in **c**. Box plots: centre line, median; box limits, 25th and 75th centiles; whiskers, minimum and maxiumum; dots, outliers. $^{*}P = 0.0356$ (**d**), $P = 0.8587$ (**e**; NS, not significant); two-tailed non-parametric Wilcoxon rank-sum test. Pre- or post-replicative chromatin assigned based on $MCM^{+}$ or $MCM^{-}$ status. **f**, Schematic depiction of 53BP1-MD. **g**, Western blot of U2OS cells with endogenously tagged 53BP1–GFP immunostained for 53BP1, GFP and loading control (MCM2). **h**, Junction PCR showing homozygous 53BP1 tagging. **i**–**k**, 3D-SIM of 53BP1 MDs in endogenously tagged U2OS-53BP1–GFP cells (**i**) or U2OS cells immunostained with mouse (**j**) or rabbit (**k**) 53BP1 antibodies, exposed to irradiation (1 Gy, 2 h). Scale bars, 100 nm (**a**); 200 nm (**c**, **i**–**k**). Experiments in **b**, **d**, **e**, **g**–**k** were biologically replicated twice with similar results. For detailed image information see Supplementary Table 1. For gel source data see Supplementary Fig. 1.



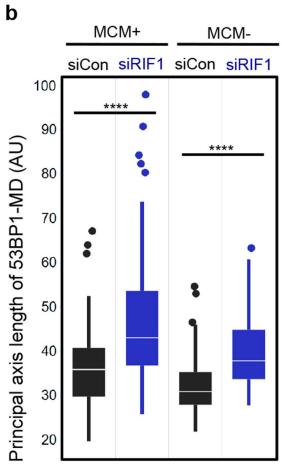

**Extended Data Fig. 2** | See next page for caption.

**Extended Data Fig. 2 | 53BP1-MD relation to underlying chromatin. a**, 3D-SIM of GFP–53BP1-MDs in U2OS cells exposed to irradiation (1 Gy, 2 h) and immunostained for γH2AX (PCC = 0.93, $n$ = 300 MDs) shows high colocalization of 53BP1 and γH2AX. **b**, STED of a γH2AX-MD in U2OS cells treated as in **a**. **c**, 3D-SIM of three different $z$-planes of HeLa cells expressing histone H2B–GFP, treated with 10 ng ml$^{-1}$ NCS for 2 h and immunostained for γH2AX. Nuclear DNA was visualized by DAPI. Insets are magnified γH2AX-MDs. Intensity line profiles of the three fluorophores (along the white line in the insets) show colocalization of chromatin with γH2AX-MDs. **d**, Western blotting of U2OS cells treated with control or *XRCC4* siRNA immunostained for XRCC4 and loading marker (KAP1). **e**, **f**, Intensity line profiles of 53BP1-MDs with XRCC4 (**e**) and RPA (**f**) in cells treated as in Fig. 1b; six independent examples per condition are shown. Fluorescence intensities in **c**, **e**, **f** were normalized to the maximum value of each profile. Scale bars, 200 nm in **a**, **b**, insets in **c** and 5 µm in whole-nucleus images (**c**). Experiments were biologically replicated twice with similar results. For detailed image information see Supplementary Table 1. For gel source data see Supplementary Fig. 1.

a



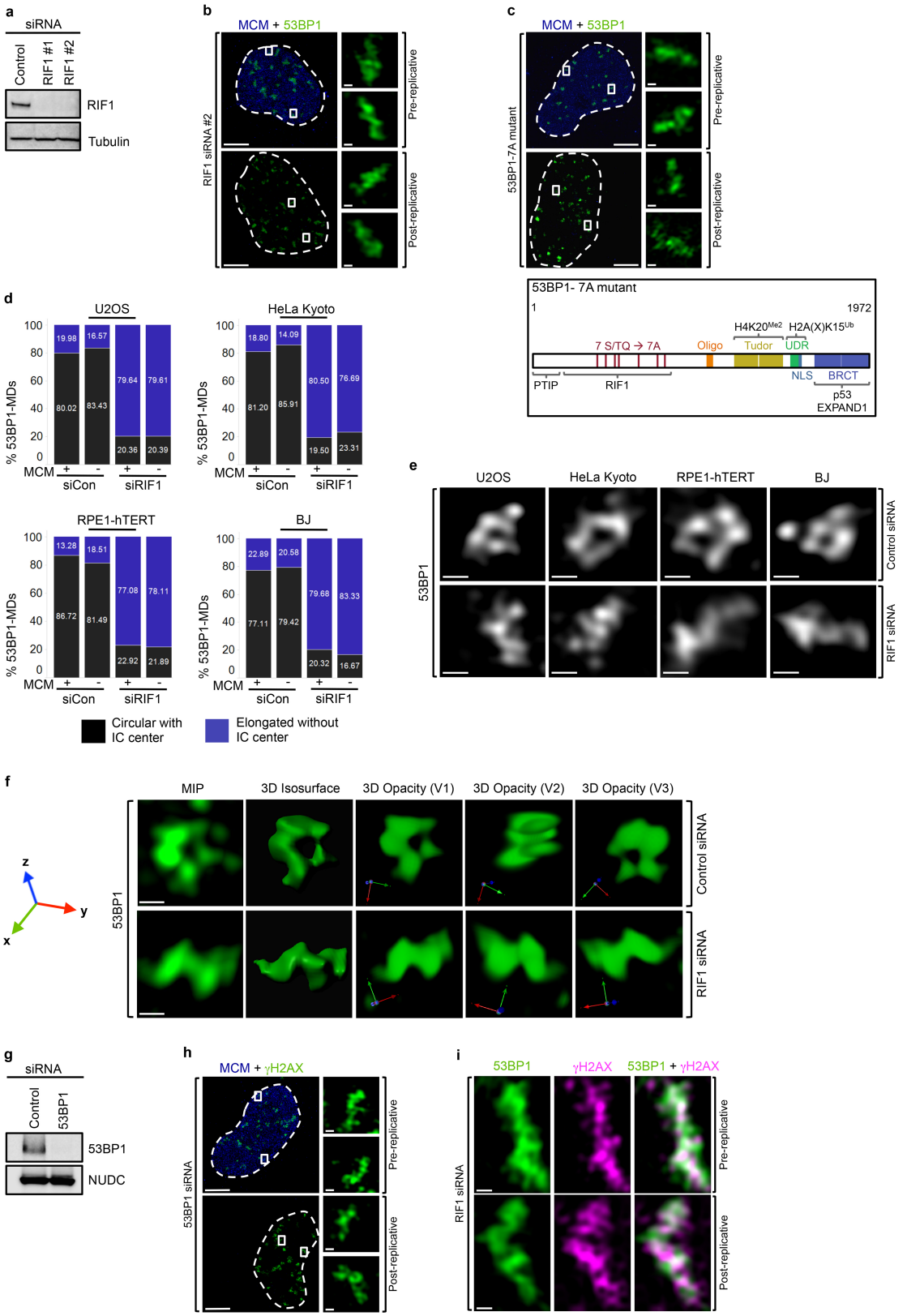

b




c

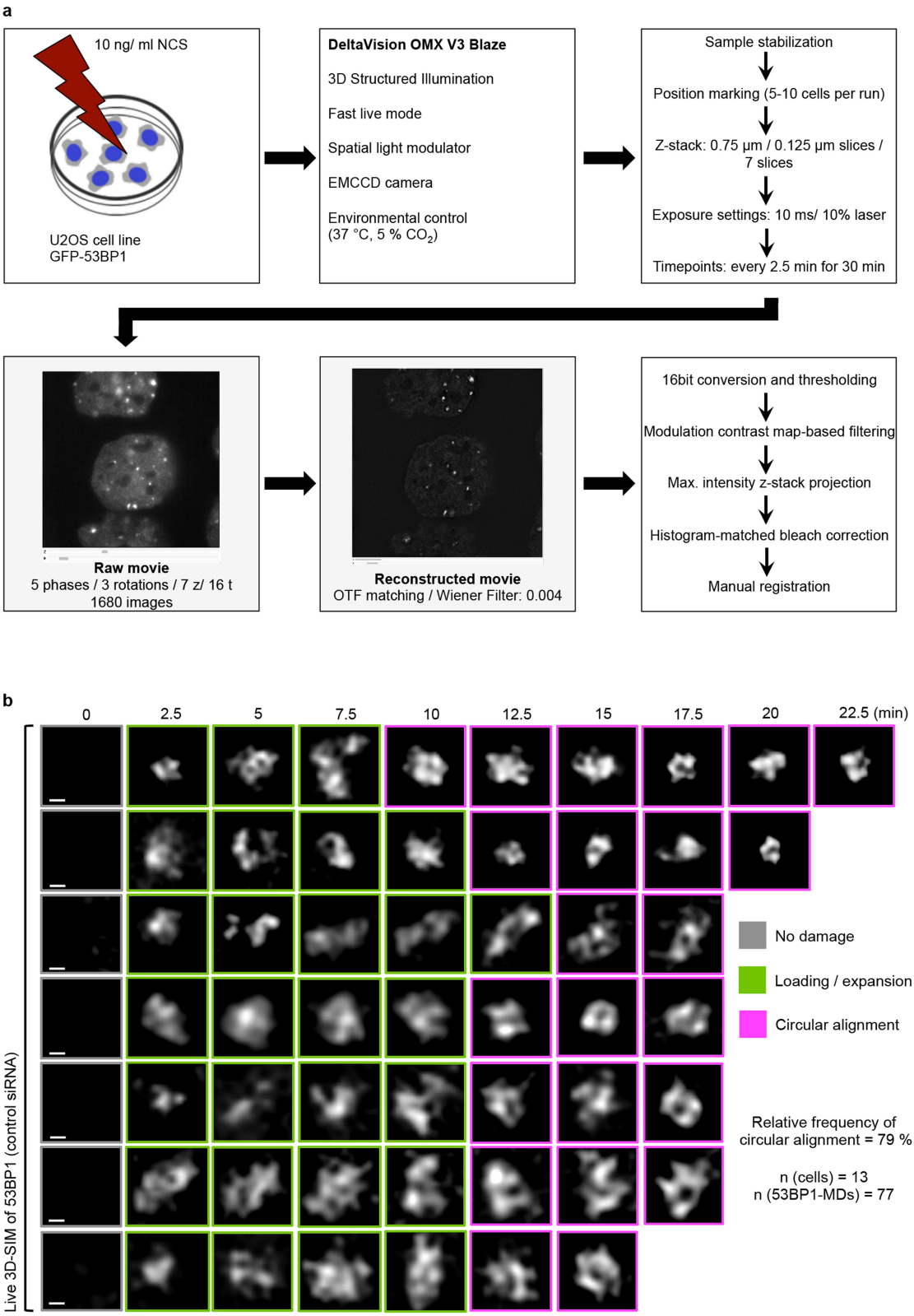
| Features | Mean breadth | Principal axis length |
|---|---|---|
| Mean breadth | 1 | 0.59 ($p = 2.74 \times 10^{-35}$) |
| Principal axis length | 0.59 ($p = 2.74 \times 10^{-35}$) | 1 |

**Extended Data Fig. 3 | Image analysis software QUANTEX and feature comparison for maximum linear dimension. a**, QUANTEX 3D image analysis workflow to analyse spatial features of 53BP1-MDs at sites of DNA damage. Step 1: 3D-SIM images were processed and segmented for cell nuclei and 53BP1-MDs using a slice-by-slice segmentation approach. Step 2: measurements of texture, morphology and geometry features were automatically derived for all segmented structures and 3D models for visual inspection were generated. Step 3: Data analysis and statistics. For more information, see Methods. **b**, QUANTEX analysis of principal axis length metric of 53BP1-MDs in cells treated with control or *RIF1* siRNAs. Principal axis length data were derived from the same experiments as in Fig. 1a, d and represent a parallel data analysis to the metric mean breadth in Fig. 1f; $n = 60$. Box plots: centre lines, medians; box limits, 25th and 75th centiles; whiskers, minimum and maximum; dots, outliers. ****$P = 9.4329 \times 10^{-6}$ (left), $2.3092 \times 10^{-9}$ (right); two-tailed non-parametric Wilcoxon rank-sum test. The experiment was biologically replicated twice with similar results. **c**, Spearman's Correlation $R^2$ value was calculated for mean breadth and principal axis length metrics derived from control (negative class, $n = 90$) and RIF1 depletion (positive class, $n = 87$) experiments combined, in order to test association. ****$P = 2.74 \times 10^{-35}$; two-sided Spearman's rank correlation coefficient method.

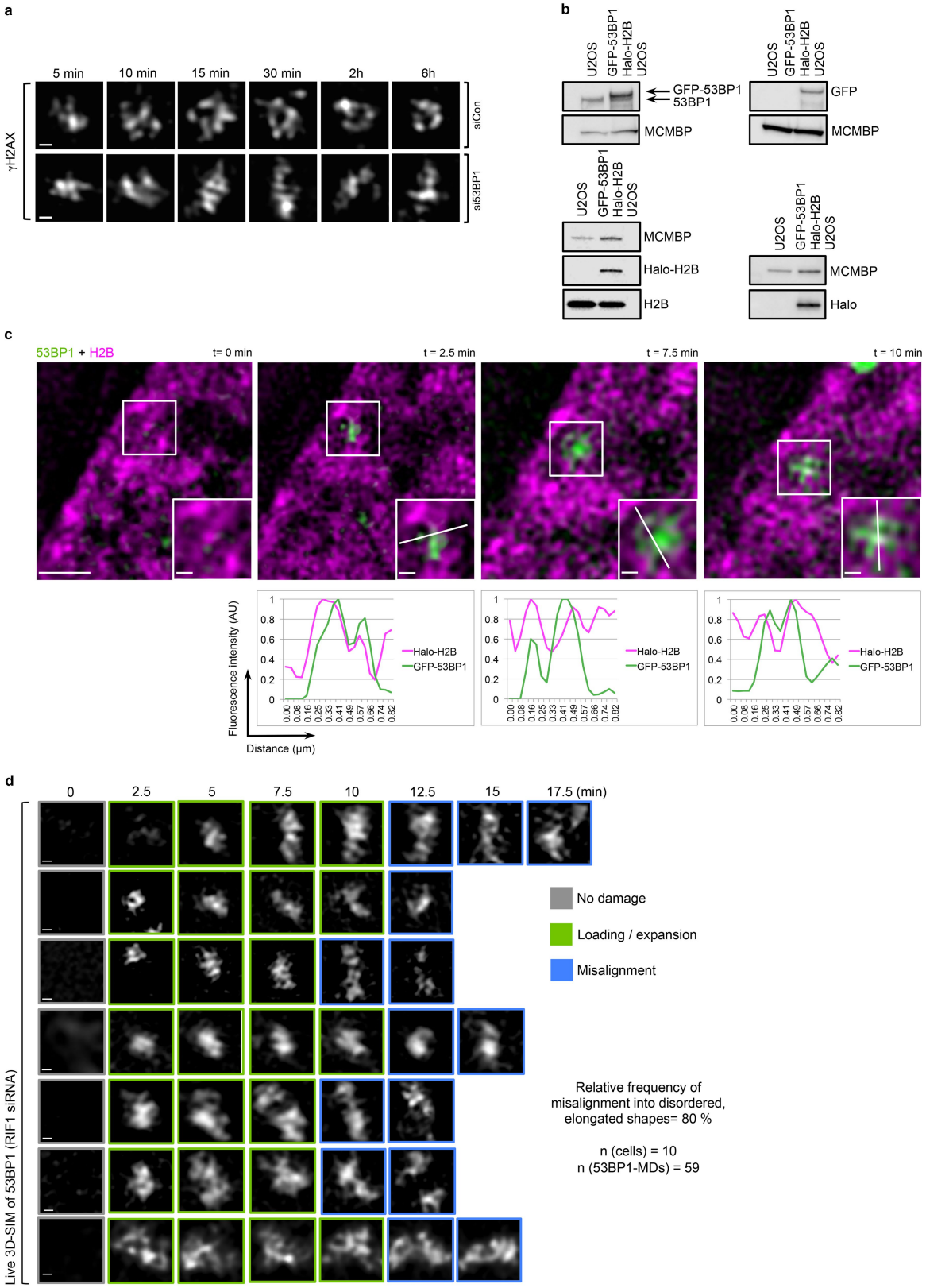



**Extended Data Fig. 4** | See next page for caption.

**Extended Data Fig. 4 | Disruption of ordered, circular arrangement of DSB-flanking chromatin after depletion of RIF1 or 53BP1. a**, Western blotting of U2OS cells treated with control or two *RIF1* siRNAs immunostained for RIF1 and loading marker (tubulin). **b**, 3D-SIM of GFP–53BP1-MDs in U2OS cells transfected with *RIF1* siRNA #2 and treated as in Fig. 1d. **c**, 3D-SIM of 53BP1-MDs in U2OS cells expressing siRNA-resistant GFP–53BP1-7A mutant and depleted for endogenous 53BP1 (#2 siRNA), exposed to irradiation (1 Gy, 2 h) (top). A schematic depiction of 53BP1-7A where glutamines in 7 SQ/TQ sites are converted to alanines (bottom). **d**, Distribution of circular with central interchromatin space (IC centre) versus elongated (no IC centre) 53BP1-MDs in U2OS, HeLa Kyoto, RPE1-hTERT and BJ cells ($n = 130$ per condition) in control or RIF1-depleted cells treated with irradiation (1 Gy, 2 h). **e**, 3D-SIM of immunostained 53BP1-MDs in U2OS, HeLa Kyoto, RPE1-hTERT and BJ cells after control or RIF1 depletion and irradiation exposure (1 Gy, 2 h). **f**, A representative 3D view of an ordered, circular arrangement of GFP–53BP1-NDs in wild-type conditions (top) and disordered, elongated shapes after RIF1 depletion (bottom). MIP, maximal intensity projection; 3D opacity view is displayed in three orientations (V1–3) indicated by coloured arrows. All 3D-SIM images in this study were routinely inspected in this way. **g**, Western blotting of U2OS cells treated with *TP53BP1* siRNA and immunostained for 53BP1 and loading marker (NUDC). **h**, 3D-SIM of γH2AX-MDs in U2OS cells transfected with *TP53BP1* siRNA and exposed to irradiation (1 Gy, 2 h). **i**, 3D-SIM of GFP–53BP1-MD in U2OS cells immunostained for γH2AX and treated as in Fig. 1d. Insets in **b**, **c**, **h** represent magnified single 53BP1-MDs. Scale bars, 5 µm in whole-nucleus images (**b**, **c**, **h**), 200 nm in **e**, **f**, **i** and insets (**b**, **c**, **h**). Experiments were biologically replicated twice with similar results. For detailed image information see Supplementary Table 1. For gel source data see Supplementary Fig. 1.

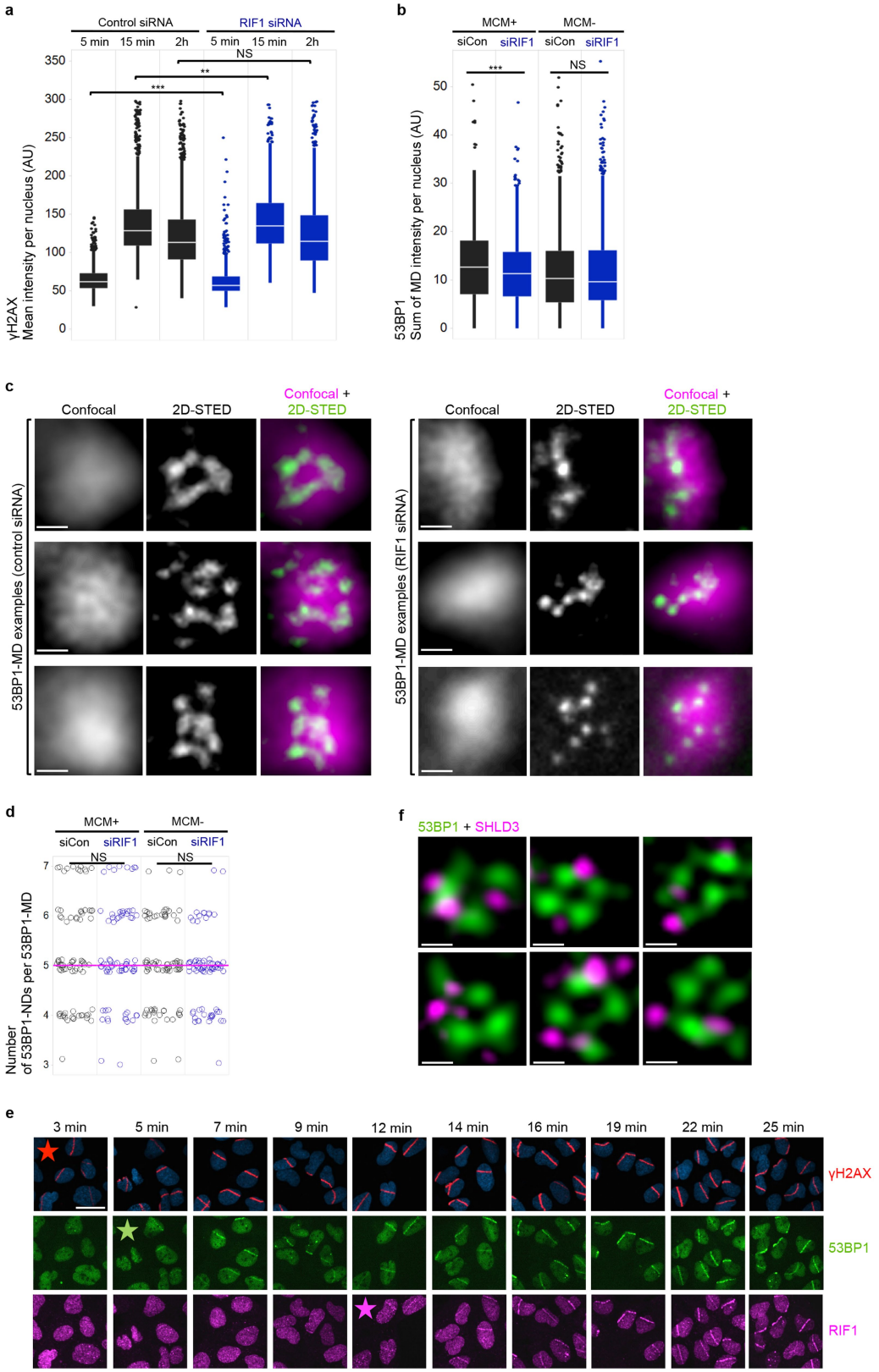



**Extended Data Fig. 5 | Live-SIM imaging of 53BP1-MDs; workflow and dynamics in control cells. a**, Schematic depiction of live-cell 3D-SIM workflow. **b**, Live-SIM of a chromosome locus harbouring DNA breakage under wild-type conditions. U2OS–GFP-53BP1 cells were treated with 10 ng ml$^{-1}$ NCS to induce DSBs and imaged immediately for up to 22.5 min at 2.5-min intervals. Image galleries for seven fields from four independent acquisitions are displayed. Manual classification of transition stages is colour-coded. Scale bars, 200 nm. For detailed image information see Supplementary Table 1.



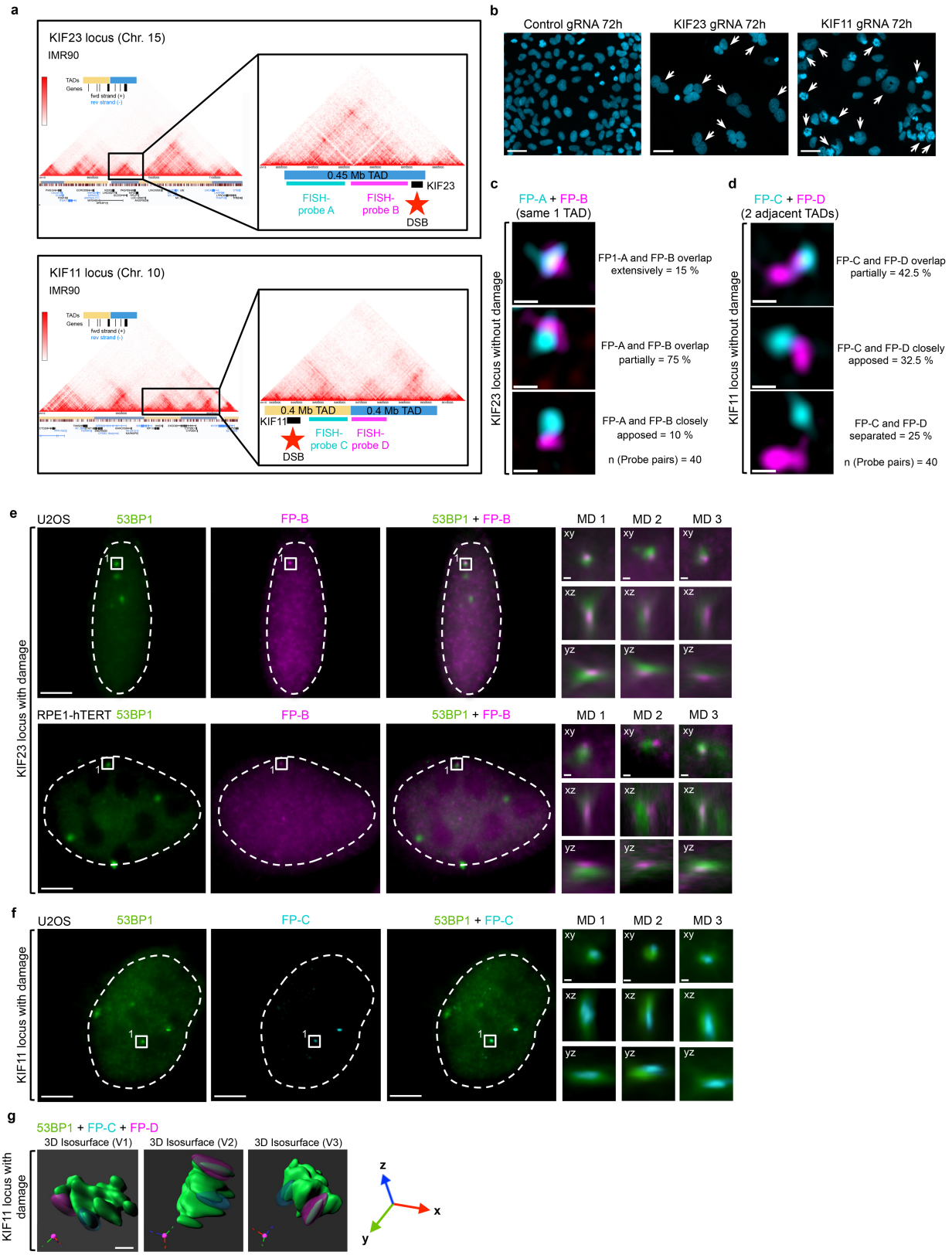

**Extended Data Fig. 6** | See next page for caption.

**Extended Data Fig. 6 | Live-SIM imaging of 53BP1-MDs with the underlying chromatin and after RIF1 depletion. a**, 3D-SIM of immunostained γH2AX-MDs in control or 53BP1-depleted U2OS cells treated with irradiation (1 Gy) for the indicated times. **b**, Western blotting of U2OS cells expressing GFP–53BP1 and H2B–Halo-Tag immunostained for 53BP1, GFP, H2B, Halo-Tag and loading marker (MCMBP). **c**, Live-SIM depicting an evolving GFP–53BP1-MD at a single H2B–HaloTag-labelled chromatin locus after induction of DSBs by NCS (10 ng ml$^{-1}$) for the indicated time-points. Insets are magnified 53BP1-MDs. Intensity line profiles of the two fluorophores (along the white lines in the insets) show colocalization of underlying chromatin with the 53BP1-MD. Fluorescence intensities were normalized to the maximum value of each profile. **d**, Additional examples of live-SIM of cells treated as in Fig. 2c. Image galleries for seven fields from four independent acquisitions are displayed. Manual classification of transition stages is colour-coded. Experiments in **a**–**c** were biologically replicated twice with similar results. Scale bars, 200 nm in **a**, **d**, and insets in **c**; 1 µm in large fields in **c**. For detailed image information see Supplementary Table 1. For gel source data see Supplementary Fig. 1.



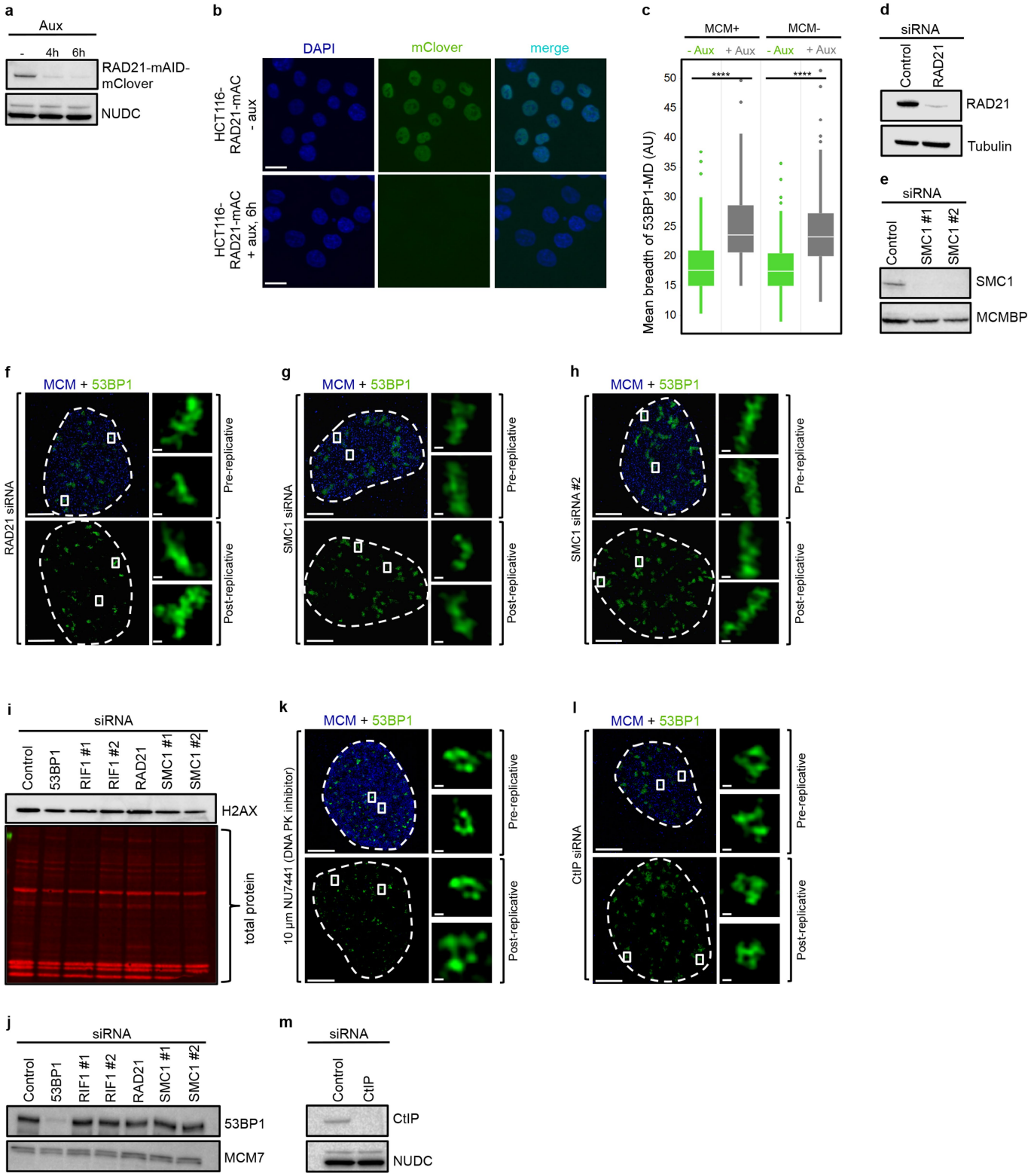

**Extended Data Fig. 7** | See next page for caption.

**Extended Data Fig. 7 | Analysis of RIF1 depletion, shieldin localization, and RIF1 recruitment dynamics in the context of DSB-flanking chromatin.** **a**, **b**, QIBC of fluorescence intensities associated with γH2AX-MDs (**a**; $n = 1{,}000$ cells per condition) and 53BP1-MDs (**b**; $n = 1{,}800$ cells per condition) in control or RIF1-depleted cells treated with irradiation (1 Gy) as indicated. Box plots: centre line, median; box limits, 25th and 75th centiles; whiskers, minimum and maximum; dots, outliers. $^{***}P = 2.0631 \times 10^{-10}$, $^{**}P = 4.8803 \times 10^{-04}$, NS $P = 0.8651$ (**a**, left to right); $^{***}P = 3.887 \times 10^{-9}$, NS $P = 0.7172$ (**b**, left to right); two-tailed non-parametric Wilcoxon rank-sum test. **c**, Confocal and STED acquisitions of immunostained 53BP1-MDs in U2OS cells treated with control or *RIF1* siRNAs, exposed to irradiation (1 Gy, 2 h) and displayed as single and overlay images. **d**, Counts of 53BP1-NDs per 53BP1-MD quantified from STED images in **c** ($n = 70$ per condition); horizonal bar shows median, $P = 0.2711$ (left), 0.9566 (right); Cochran–Armitage chi-square test. **e**, U2OS cells expressing endogenously tagged 53BP1–GFP were treated by laser microirradiation and immunostained for γH2AX and RIF1. Stars indicate times when γH2AX, 53BP1 and RIF1 were first detected at DSBs. **f**, 3D-SIM of 53BP1-MD and 3×-Flag-SHLD3 in U2OS cells exposed to irradiation (1 Gy, 2h) and immunostained for 53BP1 and Flag-tag (six independent examples are shown). Scale bars, 200 nm (**c**, **f**), 20 μm (**e**). Experiments were biologically replicated twice with similar results. For detailed image information see Supplementary Table 1. For gel source data see Supplementary Fig. 1.



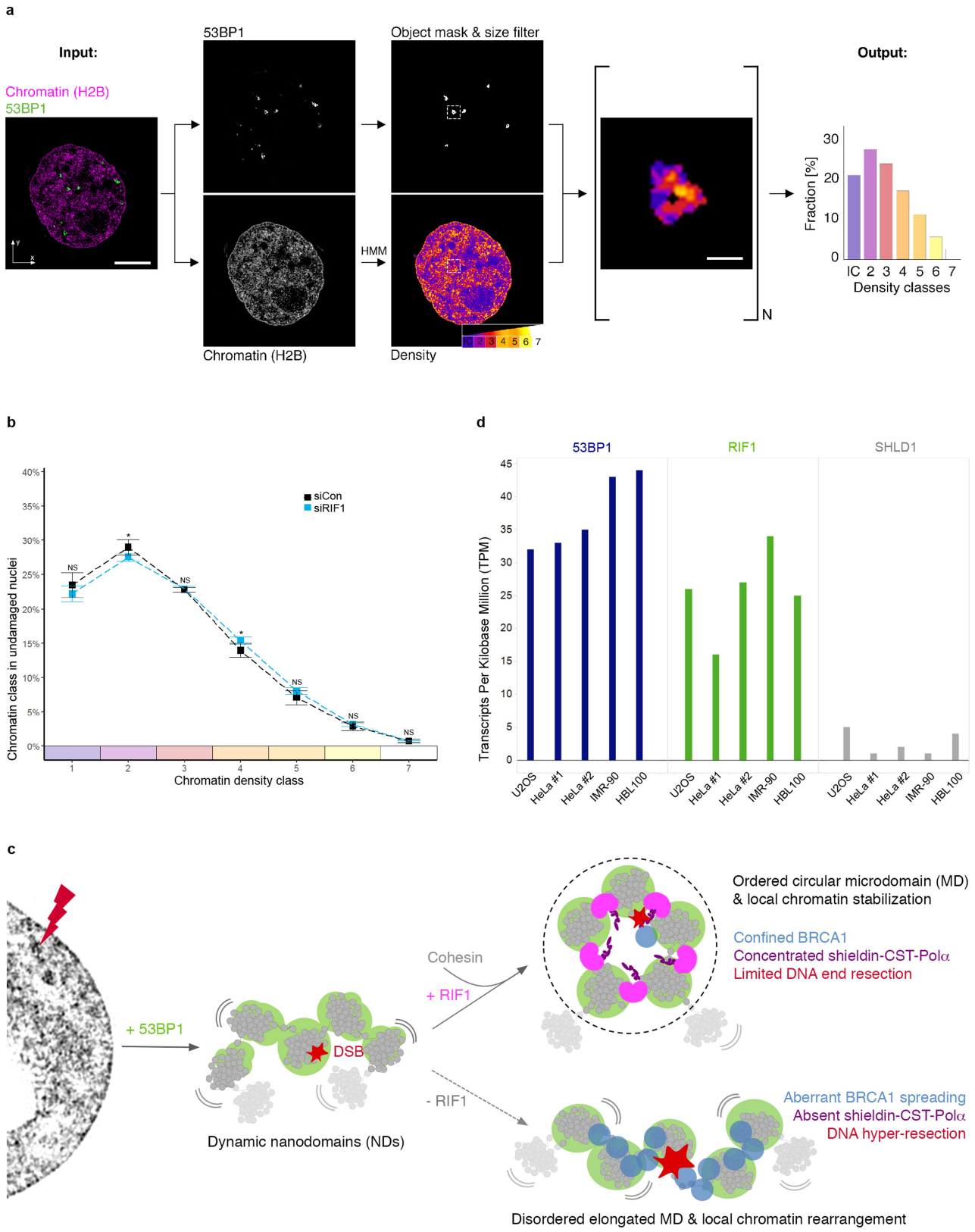

**Extended Data Fig. 8** | See next page for caption.

**Extended Data Fig. 8 | RASER-FISH analysis of 53BP1-MDs at site-specific DSBs in *KIF23* and *KIF11* loci. a**, Depiction of a 0.45-Mb TAD from a reference cell line (adapted from Yue laboratory 3D genome browser, see Methods) harbouring the *KIF23* gene (top) and a 0.4-Mb TAD harbouring the *KIF11* gene (bottom). Sites of CRISPR–Cas9 site-specific DSBs and a position of each RASER-FISH probe are indicated. **b**, DAPI-stained U2OS cells transfected with Cas9 ribonucleoprotein complexes with control, *KIF23*-, or *KIF11*-targeting guide RNAs (gRNA). Arrows indicate examples of mitotic aberrations inflicted by *KIF23* and *KIF11* knockout. **c**, **d**, 3D-SIM of the *KIF23*-TAD (**c**) and the *KIF11*-TAD (**d**) RASER-FISH probes in cells treated as in Fig. 3a, b but at loci without DNA damage (no 53BP1 signal). Dual-colour FISH probes FP-A and FP-B are located within the same TAD in **c**; FP-C and FP-D in in two adjacent TADs in **d**. **e**, Widefield microscopy of immunostained 53BP1-MDs at the damaged *KIF23*-TAD locus labelled by FP-B in U2OS and RPE1-hTERT cells 3 h after transfection with *KIF23* gRNA–Cas9. Insets (MD1–3) are magnified 53BP1-MDs shown in *xy*, *xz* and *yz* orientations. **f**, Widefield microscopy of immunostained 53BP1-MDs at the damaged *KIF11*-TAD locus labelled by FP-C in U2OS cells 3 h after transfection with *KIF11* gRNA–Cas9. Insets (MD1–3) were generated as in **e**. **g**, 3D-isosurface projections (V1–3) of 3D-SIM images of FP-C- and FP-D-labelled *KIF11* TADs after DNA damage induction as shown in Fig. 3b. Scale bars, 5 µm in whole-nucleus images (**e**, **f**), 200 nm in insets (**e**, **f**) and in **c**, **d**, 20 µm in **b**. Experiments in **b**–**f** were biologically replicated twice with similar results. For detailed image information see Supplementary Table 1.




**Extended Data Fig. 9** | See next page for caption.

**Extended Data Fig. 9 | Disruption of ordered, circular arrangement of DSB-flanking chromatin after cohesin depletion. a**, Western blotting of HCT116-RAD21–mAID–mClover cells treated with auxin (aux) as indicated and immunostained for RAD21 and loading marker (NUDC). **b**, Widefield images of HCT116-RAD21–mAID–mClover cells, either untreated or treated with auxin for 6 h to induce RAD21 degradation. **c**, QUANTEX analysis of mean breadth of 53BP1-MDs in cells treated as in Fig. 3c, d ($n = 110$). Box plots: centre line, median; box limits, 25th and 75th centiles; whiskers, minimum and maximum; dots, outliers. $^{****}P = 3.8495 \times 10^{-17}$ for $MCM^+$, $^{****}P = 7.636 \times 10^{-16}$ for $MCM^-$; two-tailed non-parametric Wilcoxon rank-sum test. **d**, Western blotting of U2OS cells treated with control or *RAD21* siRNA, immunostained for RAD21 and loading marker (tubulin). **e**, Western blotting of U2OS cells treated with control or *SMC1* siRNA, immunostained for SMC1 and loading marker (MCMBP). **f**–**h**, 3D-SIM of GFP–53BP1-MDs in U2OS cells transfected with *RAD21* siRNA (**f**), *SMC1* siRNA #1 (**g**), or *SMC1* siRNA #2 (**h**) and exposed to irradiation (1 Gy, 2 h). **i**, Western blotting of U2OS cells treated with the indicated siRNAs and immunostained for γH2AX; total protein stain is loading control. **j**, Western blotting of U2OS cells treated with indicated siRNAs and immunostained for 53BP1 and loading marker (MCM7). **k**, **l**, 3D-SIM of GFP–53BP1-MDs in U2OS cells treated with 10 μM DNA-PK inhibitor (**k**) or *RBBP8* (also known as *CtIP*) siRNA (**l**) and exposed to irradiation (1 Gy, 2 h). **m**, Western blotting of U2OS cells treated with control or *RBBP8* siRNA, immunostained for CtIP and loading marker (NUDC). Insets in (**f**–**h**, **k**, **l**) are magnified 53BP1-MDs. Scale bars, 5 μm in whole nuclei (**f**–**h**, **k**, **l**), 200 nm in insets (**f**–**h**, **k**, **l**), 20 μm in **b**. Experiments were biologically replicated twice with similar results. For detailed image information see Supplementary Table 1. For gel source data see Supplementary Fig. 1.




**Extended Data Fig. 10** | See next page for caption.

**Extended Data Fig. 10 | Chromatin density analysis by ChaiN, RNA-seq data, and a schematic model for topological surveillance of DSB loci. a**, Schematic depiction of ChaiN analysis to quantify chromatin density in 3D-SIM images based on histone H2B–GFP distribution. Reconstructed and aligned 3D-SIM images were used to segment volumes occupied by 53BP1-MDs and subjected to an HMM process to derive seven discrete GFP–H2B chromatin density classes within the segmented region. Class 1 represents chromatin-free interchromatin space, while class 2–7 feature increasing chromatin densities. An equivalent analysis of the whole nucleus serves as a control for global chromatin distributions outside 53BP1-MDs. **b**, ChaiN analysis in undamaged nuclei in wild-type or RIF1-depleted cells ($n = 12$ per condition). Median ± 95% CI. *$P = 0.0348$, 0.0226 (class 2 and 4), NS $P = 0.2525$, 0.7373, 0.0990, 0.4874, 0.9496 (classes 1, 3, 5–7); two-tailed Student's *t*-test. **c**, A hypothetical model. A DSB triggers accumulation of 53BP1 in the damaged and several neighbouring chromatin nanodomains. Saturation of 53BP1 at chromatin nanodomains prompts recruitment of RIF1 to the boundaries between them. Through functional crosstalk with cohesin, RIF1 locally stabilizes the nanodomain topology into an ordered and circular microdomain, which confines repair factors such as BRCA1 to DSBs and locally concentrates shieldin–CST–POLα to restrain DNA-end resection. Absence of RIF1 leads to topological disorder that results in excessive spreading of BRCA1, inability to concentrate DNA-end protection factors and DSB hyper-resection. **d**, RNA-seq data for *TP53BP1*, *RIF1* and *SHLD1* transcripts per million kilobases in cancerous cells (U2OS, HeLa) and normal cells (IMR90, HBL100). Data were derived from publicly available RNA-seq data in the EMBL-EBI expression atlas (see Methods). Scale bars (**a**), 5 µm in the whole nucleus and 200 nm in the magnified 53BP1-MD image (right). For detailed image information see Supplementary Table 1.

nature research

Corresponding author(s): Jiri Lukas
Nature 2018-09-12876C

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

## Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☒ | ☐ | A description of all covariates tested |
| ☒ | ☐ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☒ | ☐ | For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☒ | ☐ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒ | ☐ | Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |
| ☐ | ☒ | Clearly defined error bars<br>*State explicitly what error bars represent (e.g. SD, SE, CI)* |

*Our web collection on statistics for biologists may be useful.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | ScanR Acquisition software (Olympus, V2.7.1 for ScanR high-content microscopy)<br>Volocity acquisition software (Perkin Elmer, V6.3, for Ultraview Vox spinning disk microscopy),<br>ZEN Black acquistion software (Zeiss, for ELYRA PS.1 SIM microscopy and LSM880 confocal microscopy)<br>OMX SoftWoRx software (GE Healthcare, V6.1, for OMX V3 Blaze SIM microscopy)<br>ImSpector software (Abberior Instruments, for Abberior STED microscopy)<br>PALM-Robo software (Zeiss, 4.5.09, for laser-microirradiation microscopy) |
| Data analysis | ScanR Analysis (Olympus, V 2.7.1, for QIBC)<br>Fiji/Image J with SIMCheck plug-in (for SIM microscopy quality control)<br>QUANTEX (in house custom software developed for this study, for 3D image analysis of sub-cellular structures)<br>MATLAB (MathWorks R2018b)<br>R (3.6.1, for coding of QUANTEX software and data plotting of ChaiN data)<br>Excel (Microsoft, 2016, for image analysis data management)<br>Spotfire (Tibco, V7.8.0.1.20, for data visualization and statistics)<br>Imaris (Bitplane, V x64 9.0.2 for visualization of 3D-SIM images)<br>Volocity (Perkin Elmer, V6.3 for image viewing). |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Numerical and statistical source data for Figs. 1e,f, 2d,e, 3a,b,e,f,g,h,i and Extended Data Figs. 1d,e, 2c,e,f, 3b, 4d, 5b, 6c,d, 7a,b,d, 8c,d, 9c, 10b,d have been provided with this manuscript. Primary imaging data underlying widefield, confocal, SIM and STED images in Figs. 1a,b,c,d, 2a,b,c,f, 3a,b,c,d,e,f,g,h and Extended Data Figs. 1c,i,j,k, 2a,b,c, 4b,c,e,f,h,i, 5b, 6a,c,d, 7c,e,f, 8b,c,d,e,f,g, 9b,f,g,h,k,l has been deposited at the European Bioinformatics Institute (EBI) BioStudies database (https://www.ebi.ac.uk/biostudies/) with accession number S-BSST275. Processed imaging datasets underlying QIBC, QUANTEX, ChaiN and other analysis, including guidance on how to navigate datasets, are available from the corresponding author upon reasonable request.
There are no restrictions on data availability.

# Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Sample size | No statistical methods were used to predetermine sample size, as this study did not include animal models or human participants. Sample size was determined based on standards in the field and experimental experience to obtain statistical significance and reproducibility. |
| Data exclusions | All data acquired for for this study were included in the analysis with rare exceptions where images did not pass quality controls: for 3D-SIM microscopy, images that showed more than 50% of bleaching during acquisition and did not pass the SIMcheck quality control step for artefact-free image reconstruction were discarded. In 3D-SIM live cell imaging, images were discarded if the cell moved out of focus or out of field of view during image acquistion. In Quantex-based image analysis, subcellular structures of interest (e.g. 53BP1 MDs) that were touching the boundary of the image frame, and objects, that could not be faithfully segmented for further image analysis (e.g. over-segmentation of closely apposed touching structures of high intensity or under-segmentation of low-intensity objects) were not further analysed. |
| Replication | All experimental findings were reliably reproduced in multiple independent experiments as indicated in the figure legends. All attempts of replication were successful. |
| Randomization | No randomization was done, because this study does not involve animals or human participants. Samples were organized into groups based on treatments (e.g. untreated or control siRNA treated compared to target-specific siRNAs; experimental time-points). Appropriate controls were included in all experiments. |
| Blinding | There was no blinded group allocation. All data that passed quality controls (see data exclusion) were analyzed by unbiased automated image analysis or under strict internal standards for objective manual image analysis. |

# Reporting for specific materials, systems and methods

### Materials & experimental systems

| n/a | Involved in the study |
|---|---|
| ☐ | ☒ Unique biological materials |
| ☐ | ☒ Antibodies |
| ☐ | ☒ Eukaryotic cell lines |
| ☒ | ☐ Palaeontology |
| ☒ | ☐ Animals and other organisms |
| ☒ | ☐ Human research participants |

### Methods

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ ChIP-seq |
| ☒ | ☐ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |

## Unique biological materials

Policy information about availability of materials

Obtaining unique materials

Cell lines generated for this study (U2OS cells homozygously tagged with 53BP1-mEGFP; U2OS cells stably expressing EGFP-53BP1 and Histone H2B-HaloTag, U2OS cells stably expressing GFP-53BP1-7A mutant) will be made available upon reasonable request. Other cell lines that were published previously by our laboratory and used in this study can be obtained upon reasonable request.

## Antibodies

Antibodies used

Antibody, Host, Supplier, Cat. number, Lot number, Technique, Conc
53BP1, Mouse, Millipore, MAB3802, 3106536, IF, 1:750
53BP1, Rabbit, Novus Biologicals, NB100-305, A4, IF 1:750, WB 1:1000 for WB
53BP1, Rabbit, Novus Biologicals, NB100-304, F, WB 1:1000
BRCA1, Mouse, Calbiochem, O92, D00168480, IF 1:100
CtIP, Mouse, Active Motif, 61141; 03618004, clone 14-1, WB 1:250
FLAG-Tag, Mouse, Sigma-Aldrich, F1804, SLBT 7654, IF 1:300
GFP, Rabbit, Torrey Pines Biolabs, TP401, 081211, WB 1:1000
H2AX phospho-S139, Mouse, Abcam, ab22551, GR226782-1, IF 1:1000
H2AX phospho-S139, Rabbit, Cell Signaling, 9733, 8, IF 1:1000
HaloTag, Mouse, Promega, G921A, 0000261985, WB 1:1000
Histone H2B, Rabbit, Abcam, ab1790, GR310932, WB 1:2000
KAP1, Rabbit, Bethyl Laboratories, A300-274A, 3, WB 1:2000
MCM2, Mouse, Novus Biologicals, H00004171-M01, I3211-6A8, IF 1:200, WB 1:1000
MCM5 Rabbit, Abcam, ab17967, GR249182-35, IF 1:200
MCM7 Mouse, Santa Cruz, sc-9966, E2015, WB 1:1000 for WB
MCMBP, Rabbit, Novus Biologicals, NBP1-90746, A115112, WB 1:1000
NUDC, Rabbit, Sigma-Aldrich, HPA027183 , R12662, WB 1:1000 for WB
RAD21, Mouse, Millipore, 05-908, 3135877, clone 53A303, WB 1:500
RAP80, Rabbit, Bethyl Laboratories, A300-764A, 1, IF 1:400
RIF1, Rabbit, Bethyl Laboratories, A300-569A, 5, IF 1:500
RIF1 Mouse, Santa Cruz, sc-515573, C0216, IF 1:500
RIF1, Rabbit, Cell Signaling, 95558, IF 1:500, 1, WB 1:1000
RPA70, Rabbit, Abcam, ab79398, GR212113-23, IF 1:300
SMC1, Rabbit, Novus Biologicals, NBP2-67733, HJ0910, WB 1:1000 for WB
Tubulin, Mouse, Santa Cruz, SC-8035, C3012, WB 1:500
XRCC4, Rabbit, Abcam, ab213729, GR302718-2, IF 1: 100
Goat anti-mouse Alexa Fluor 488, Goat, Invitrogen, A11029, 1942237, IF 1:1000
Goat anti-rabbit Alexa Fluor 488, Goat, Invitrogen, A11034, 1937195, IF 1:1000
Goat anti-mouse Alexa Fluor 568, Goat, Invitrogen, A11031, 2026148, IF 1:1000
Goat anti-rabbit Alexa Fluor 568, Goat, Invitrogen, A11036, 1924788, IF 1:1000
Goat anti-mouse Alexa Fluor 647, Goat, Invitrogen, A21236, 1793803, IF 1:1000
Goat anti-rabbit Alexa Fluor 647, Goat, Invitrogen, A21245, 1805235, IF 1:1000
Goat anti-mouse STAR RED , Goat, Abberior, 2-0002-011-2, 26062018Hp, IF 1:1000
Goat anti-rabbit STAR RED , Goat, Abberior, 2-0012-011-9, 09072018CW/JR, IF 1:1000
Goat anti-mouse STAR 580 , Goat, Abberior, 2-0002-005-1 , IF 1:1000
Goat anti-rabbit STAR 580 , Goat, Abberior, 2-0012-005-8, 12102016HP, IF 1:1000
Goat anti-rabbit IgG (H&L) Peroxidase labelled, Goat, Vector Laboratories, PI-1000, ZE0614, WB 1:10000
Horse anti-mouse IgG (H&L) Peroxidase labelled, Horse, Vector Laboratories, PI-2000, ZC212, WB 1:10000

Validation

All antibody validations were done with cell line samples derived from human origin.
53BP1, Mouse, Millipore, MAB3802: validated by IF after 53BP1 siRNA in our laboratory
53BP1, Rabbit, Novus Biologicals, NB100-305, validated by IF after 53BP1 siRNA in our laboratory
53BP1, Rabbit, Novus Biologicals, NB100-304, validated by WB (Extended Data Fig. 4b)
BRCA1, Mouse, Calbiochem, O92, D00168480, IF 1:100, validated by IF after Brca1 siRNA in our laboratory
CtIP, Mouse, Active Motif, 61141, validated by WB after CtIP siRNA (Extended Data Fig. 9m)
FLAG-Tag, Mouse, Sigma-Aldrich, F1804, see manufacturer information (https://www.sigmaaldrich.com/catalog/product/sigma/f1804?lang=en&region=DK)
GFP, Rabbit, Torrey Pines Biolabs, TP401, validated by WB (Extended Data Fig. 1b)
H2AX phospho-S139, Mouse, Abcam, ab22551, validated by WB (Extended Data Fig. 9i) and by IF in our laboratory
H2AX phospho-S139, Rabbit, Cell Signaling, 9733, validated by IF in our laboratory
HaloTag, Mouse, Promega, G921A, validated by WB (Extended Data Fig. 6b)
Histone H2B, Rabbit, Abcam, ab1790, validated by WB (Extended Data Fig. 6b)
KAP1, Rabbit, Bethyl Laboratories, A300-274A, 3, validated by WB (Extended Data Fig. 2d)
MCM2, Mouse, Novus Biologicals, H00004171-M01, validated by IF in our laboratory and WB (Extended Data Fig. 1g))
MCM5 Rabbit, Abcam, ab17967, validated by IF in our laboratory
MCM7 Mouse, Santa Cruz, sc-9966, validated by WB (Extended Data Fig. 9l)
MCMBP, Rabbit, Novus Biologicals, NBP1-90746, validated by WB (Extended Data Figs. Fig. 6b, 9e)
NUDC, Rabbit, Sigma-Aldrich, HPA027183 , R12662, validated by Human Protein Atlas (https://www.sigmaaldrich.com/catalog/product/sigma/f1804?lang=en&region=DK)

RAD21, Mouse, Millipore, 05-908, 3135877, validated by WB (Extended Data Fig. 9d)
RAP80, Rabbit, Bethyl Laboratories, A300-764A, see manufacturer information (https://www.bethyl.com/product/A300-764A?referrer=search)
RIF1, Rabbit, Bethyl Laboratories, A300-569A, see manufacturer information (https://www.bethyl.com/product/A300-569A?referrer=search)
RIF1, Rabbit, Cell Signaling, 95558, validated by WB (Extended Data Fig. 4a)
RPA70, Rabbit, Abcam, ab79398, see manufacturer information (https://www.abcam.com/rpa70-antibody-epr3472-ab79398.html)
SMC1, Rabbit, Novus Biologicals, NBP2-67733, validated by WB (Extended Data Fig. 9e)
Tubulin, Mouse, Santa Cruz, SC-8035, see manufacturer information (https://www.scbt.com/scbt/product/alpha-tubulin-antibody-tu-02?requestFrom=search)
XRCC4, Rabbit, Abcam, ab213729, validated by WB (Extended Data Fig. 2d)

## Eukaryotic cell lines

Policy information about cell lines

| | |
|---|---|
| Cell line source(s) | RPE1-hTERT cells and BJ fibroblasts were obtained from ATCC. Parental U2OS cell line was originally obtained from Ed Harlow´s lab in Boston (1990´s) and maintained in the cell line repository of the Danish Cancer Society. U2OS cell lines expressing fluorescently tagged proteins (GFP-53BP1 WT or 53BP1-7A mutant, H2B-Halo/GFP-53BP1 or endogenously tagged 53BP1-GFP) are derivates of the parental U2OS cell line. U2OS cell stably overexpressing 3xFLAG-RINN1/Shld3 were obtained from Chuna Choudhary. HeLa-Kyoto cells were obtained from Dr. S. Narumiya. HeLa cells stably expressing Histone H2B-GFP were obtained from Francis Barr. HCT116 cells with an endogenously integrated RAD21-mAID-mClover were obtained from Erez Lieberman Aiden. |
| Authentication | All cell lines have been authenticated by STR profiling. |
| Mycoplasma contamination | All cell lines are routinely (two month basis) tested for mycoplasma using PCR-based methods (LONZA) and always found negative. |
| Commonly misidentified lines (See ICLAC register) | Cell lines used in this study were not listed in the commonly misidentified category. |

Article

# Metabolic regulation of gene expression by histone lactylation

https://doi.org/10.1038/s41586-019-1678-1

Received: 21 June 2018

Accepted: 13 September 2019

Published online: 23 October 2019

Di Zhang[1,12], Zhanyun Tang[2,12], He Huang[1,9], Guolin Zhou[1], Chang Cui[1], Yejing Weng[1], Wenchao Liu[1], Sunjoo Kim[3], Sangkyu Lee[3], Mathew Perez-Neut[1], Jun Ding[1], Daniel Czyz[4], Rong Hu[5,6], Zhen Ye[5,6], Maomao He[7], Y. George Zheng[7], Howard A. Shuman[4], Lunzhi Dai[1,10], Bing Ren[5,6], Robert G. Roeder[2], Lev Becker[1,8,11]* & Yingming Zhao[1,8]*

The Warburg effect, which originally described increased production of lactate in cancer, is associated with diverse cellular processes such as angiogenesis, hypoxia, polarization of macrophages and activation of T cells. This phenomenon is intimately linked to several diseases including neoplasia, sepsis and autoimmune diseases[1,2]. Lactate, which is converted from pyruvate in tumour cells, is widely known as an energy source and metabolic by-product. However, its non-metabolic functions in physiology and disease remain unknown. Here we show that lactate-derived lactylation of histone lysine residues serves as an epigenetic modification that directly stimulates gene transcription from chromatin. We identify 28 lactylation sites on core histones in human and mouse cells. Hypoxia and bacterial challenges induce the production of lactate by glycolysis, and this acts as a precursor that stimulates histone lactylation. Using M1 macrophages that have been exposed to bacteria as a model system, we show that histone lactylation has different temporal dynamics from acetylation. In the late phase of M1 macrophage polarization, increased histone lactylation induces homeostatic genes that are involved in wound healing, including *Arg1*. Collectively, our results suggest that an endogenous 'lactate clock' in bacterially challenged M1 macrophages turns on gene expression to promote homeostasis. Histone lactylation thus represents an opportunity to improve our understanding of the functions of lactate and its role in diverse pathophysiological conditions, including infection and cancer.

Inspired by the discovery of various histone acylations derived from cellular metabolites[3,4], we predicted and identified lysine lactylation (Kla) as a new type of histone mark that can be stimulated by lactate (Fig. 1a). Initial evidence for histone Kla came from the observation of a mass shift of 72.021 Da on lysine residues in three proteolytic peptides that were detected in high-performance liquid chromatography (HPLC)–tandem mass spectrometry (MS/MS) analysis of tryptically digested core histones from human MCF-7 cells (Fig. 1b and Extended Data Fig. 1b, d). This mass shift is the same as that caused by the addition of a lactyl group to the ε-amino group of a lysine residue.

To validate the existence of lysine lactylation in histones, we used four orthogonal methods[5]. In the first two methods, we used HPLC–MS/MS analysis to compare a synthetic peptide with its in vivo-derived counterpart to determine whether the two versions of the peptide have the same chemical properties in terms of chromatographic elution in HPLC analysis and fragmentation pattern in MS/MS analysis. To achieve this, we generated three histone peptides bearing Kla modifications: H3K23-QLATK$_{la}$AAR; H2BK5-PELAK$_{la}$SAPAPK; and H4K8-GGK$_{la}$GLGK. Each pair of peptides co-eluted in HPLC and had comparable MS/MS spectra (Fig. 1b and Extended Data Fig. 1a–e). To confirm the modification further, we developed a pan anti-Kla antibody (Extended Data Fig. 1f, g). Immunoblots using the pan anti-Kla antibody confirmed the presence of histone Kla and showed that histone Kla levels were increased in a dose-dependent fashion in response to exogenous L-lactate (Extended Data Fig. 1h–j). Subsequent mass spectrometry analyses identified 26 and 16 histone Kla sites from human HeLa cells and mouse bone marrow-derived macrophages (BMDMs), respectively (Fig. 1c). Finally, metabolic labelling experiments using isotopic sodium L-lactate ($^{13}C_3$) followed by MS/MS analysis demonstrated that lysine lactylation can be derived from lactate (Extended Data Fig. 1k). Together, these experiments demonstrate that histone Kla is an in vivo protein post-translational modification derived from lactate.

[1]Ben May Department for Cancer Research, The University of Chicago, Chicago, IL, USA. [2]Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, New York, NY, USA. [3]BK21 Plus KNU Multi-Omics based Creative Drug Research Team, College of Pharmacy, Research Institute of Pharmaceutical Sciences, Kyungpook National University, Daegu, South Korea. [4]Department of Microbiology, The University of Chicago, Chicago, IL, USA. [5]Ludwig Institute for Cancer Research, University of California at San Diego, La Jolla, CA, USA. [6]Center for Epigenomics and Department of Cellular and Molecular Medicine, University of California, San Diego School of Medicine, La Jolla, CA, USA. [7]Department of Pharmaceutical and Biomedical Sciences, University of Georgia, Athens, GA, USA. [8]University of Chicago Medicine Comprehensive Cancer Center, Chicago, IL, USA. [9]Present address: Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, China. [10]Present address: Department of General Practice, State Key Laboratory of Biotherapy, West China Hospital, Sichuan University, and Collaborative Innovation Center of Biotherapy, Chengdu, China. [11]Ludwig Center for Metastasis Research, The University of Chicago, Chicago, IL, USA. [12]These authors contributed equally: Di Zhang, Zhanyun Tang. *e-mail: levb@uchicago.edu; yingming.zhao@uchicago.edu

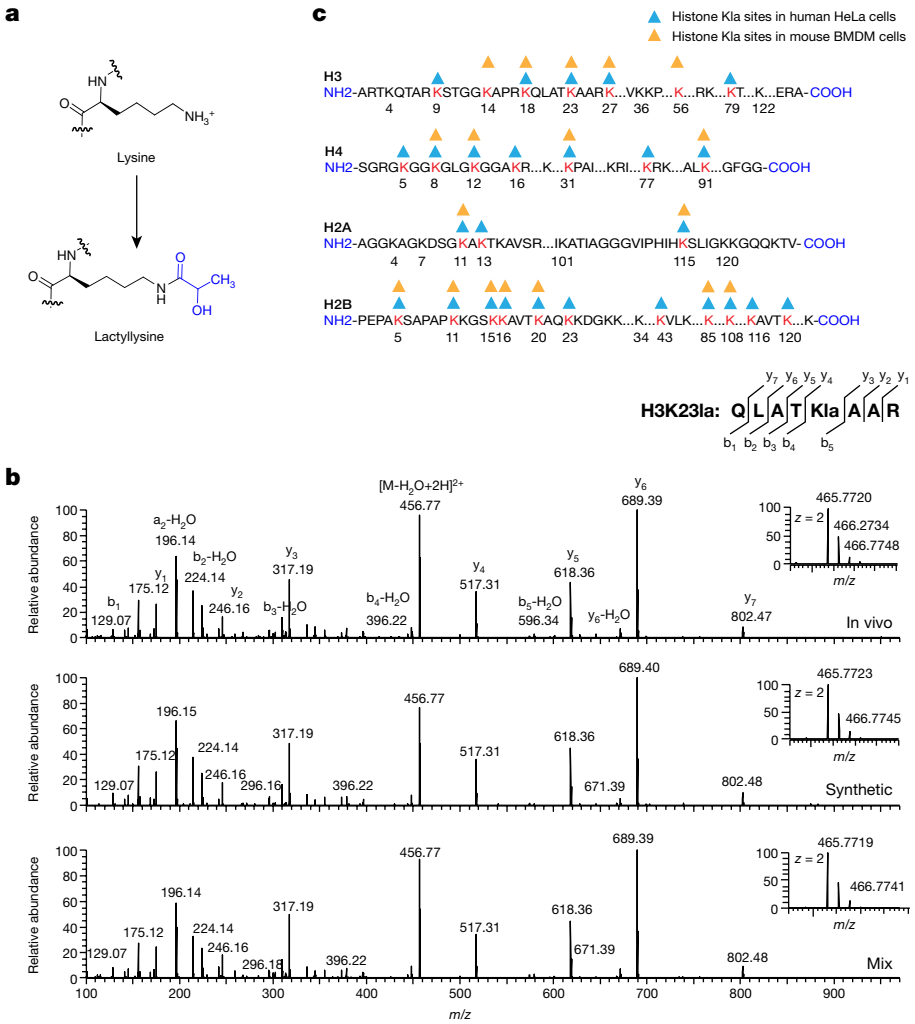



**Fig. 1 | Identification and validation of histone Kla. a**, Illustration of Kla structure. **b**, MS/MS spectra of a lactylated histone peptide (H3K23la) derived from MCF-7 cells (in vivo), its synthetic counterpart, and their mixture. The b ion refers to the N-terminal parts of the peptide, and the y ion refers to the C-terminal parts of the peptide. Data represent two independent experiments. **c**, Illustration of histone Kla sites identified in human and mouse cells.

## Histone Kla is regulated by glycolysis

Given that extracellular lactate can stimulate histone Kla, we hypothesized that modulation of intracellular lactate production would also affect histone Kla levels. We exposed MCF-7 and other cell lines to various concentrations of glucose, the major source of intracellular lactate. Both lactate production and histone Kla levels were induced by glucose in a dose-dependent manner (Fig. 2a, b and Extended Data Fig. 2a–c). Conversely, the non-metabolizable glucose analogue 2-deoxy-D-glucose (2-DG) decreased both lactate production and histone Kla levels (Fig. 2c, d). Furthermore, metabolic labelling experiments using isotopic glucose (U-$^{13}C_6$) followed by MS/MS analysis demonstrated that lysine lactylation is endogenously derived from glucose (Extended Data Fig. 2d and Supplementary Table 1). Quantitative proteomics analysis across a diverse set of histone sites demonstrated that histone Kla and Kac have different kinetics of U-$^{13}C_6$-glucose incorporation in MCF-7 cells (Extended Data Fig. 2e, f). $^{13}$C-labelled histone Kac reached a steady state at 6 h—similar to the previous observation in HCT116 cells[6]. By contrast, histone Kla increased over a 24-h time course (Extended Data Fig. 2e, f). Immunoblotting results corroborated the MS/MS data in cell lines such as MCF-7 cells (Extended Data Fig. 2i–k).

Lactate production is determined by the balance between glycolysis and mitochondrial metabolism. We tested whether the activities of enzymes in these two pathways can modulate lactate levels that in turn regulate histone Kla (Fig. 2e). Sodium dichloroacetate (DCA) and oxamate were used to inhibit lactate production by modulating the activities of pyruvate dehydrogenase and lactate dehydrogenase, respectively. As anticipated, intracellular levels of lactate were decreased by these two compounds (Fig. 2f) and levels of histone Kla were lowered (Fig. 2g, h). By contrast, rotenone—an inhibitor of the mitochondrial respiratory chain complex I that drives cells towards glycolysis—increased levels of both intracellular lactate and histone Kla (Fig. 2f, i). Quantification of histone Kla and Kac marks by stable isotope labelling with amino acids in cell culture (SILAC) and MS/MS analyses corroborated the immunoblot data from DCA- and rotenone-treated MCF-7 cells (Extended Data Fig. 2l, m). Furthermore, labelling experiments with U-$^{13}C_6$-glucose showed that the incorporation of $^{13}$C into histone Kla but not Kac was decreased by DCA (Extended Data Fig. 2e–h). Together, these observations demonstrate that endogenous production of lactate is a key determinant of histone Kla levels.

## Hypoxia and bacterial exposure stimulate histone Kla

Increased glycolysis and lactate production are coupled with diverse cellular processes. To investigate whether histone Kla is regulated by glycolysis under physiological conditions, we chose two model systems: hypoxia and M1 macrophage polarization. In response to hypoxia, cells reprogram their metabolism by inhibiting oxidative phosphorylation

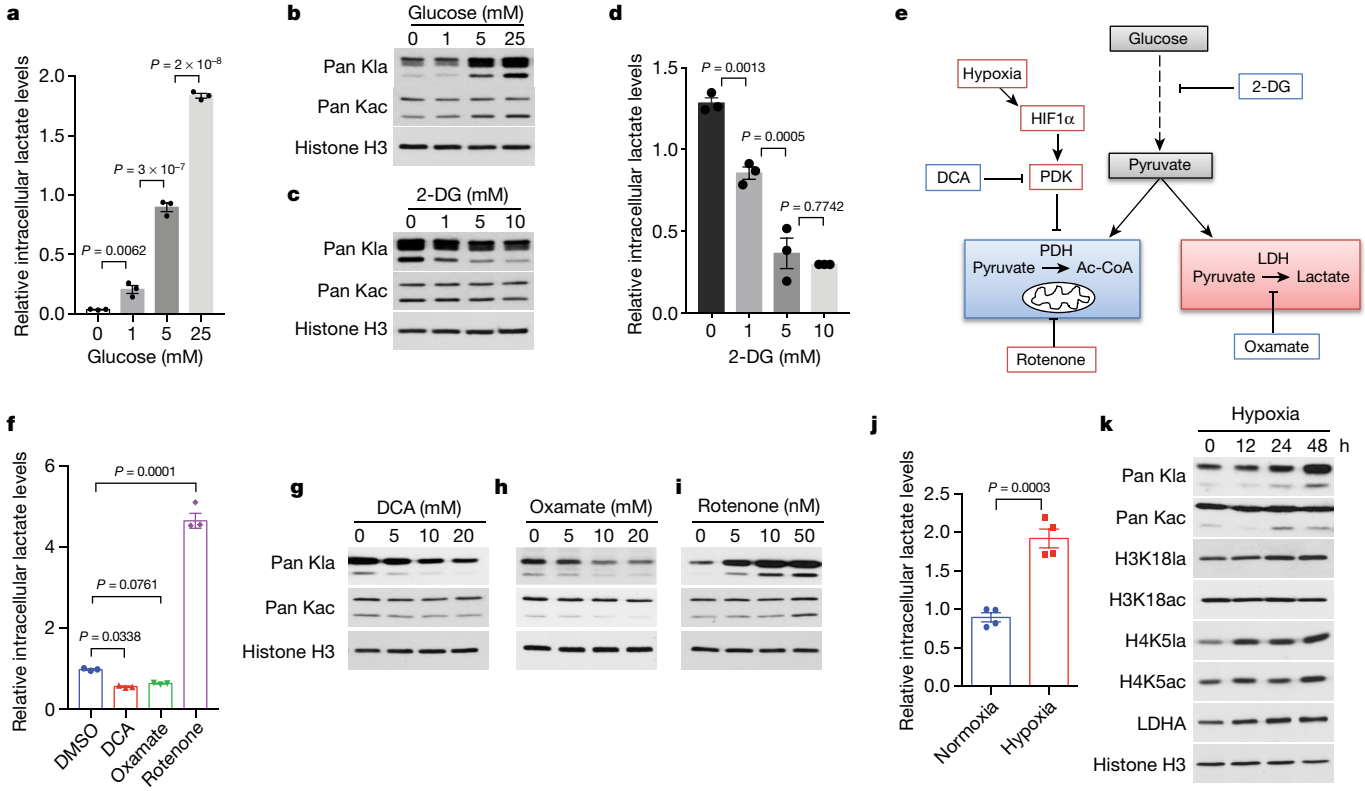



**Fig. 2 | Lactate regulates histone Kla. a–d**, Intracellular lactate levels (**a**, **d**) and histone Kla levels (**b**, **c**) were measured from MCF-7 cells cultured in different concentrations of glucose or 2-DG in the presence of 25 mM glucose for 24 h. Lactate was measured by a lactate colorimetric kit. $n = 3$ biological replicates; statistical significance was determined using one-way ANOVA followed by Sidak's multiple comparisons test. Immunoblots were performed using acid-extracted histone samples. The pan anti-Kla and anti-Kac immunoblots indicate molecular masses between 10 and 15 kDa. **e**, Regulation of glycolysis and lactate production by diverse metabolic modulators. **f**, Intracellular lactate levels were measured in MCF-7 cells treated with indicated glycolysis modulators for 24 h. $n = 3$ biological replicates; statistical significance was determined using one-way ANOVA followed by Dunnett's multiple comparisons test. **g–i**, Immunoblots of acid-extracted histones (rotenone and DCA) or whole-cell lysates (oxamate) from MCF-7 cells in response to different glycolysis modulators. **j**, Intracellular lactate levels were measured in MCF-7 cells in response to hypoxia. $n = 4$ biological replicates; statistical significance was determined using unpaired two-tailed *t*-test. **k**, Immunoblots of acid-extracted histones from MCF-7 cells under hypoxia (1% oxygen) for indicated time points. Data in **a**, **d**, **f** and **j** are mean and s.e.m. Data in **b**, **c**, **g–i** and **k** represent three independent experiments.

and enhancing glycolysis, which stimulates the production of lactate[7]. Hypoxia induced intracellular production of lactate and increased histone Kla levels but not Kac levels in MCF-7 cells (Fig. 2j, k and Extended Data Fig. 3a–d). SILAC-based mass spectrometric quantification of histone Kla and Kac confirmed the immunoblotting data (Extended Data Fig. 3e, f). Similar results were obtained in HeLa and RAW 264.7 cells (Extended Data Fig. 3g, h). Furthermore, we found that the induction of lactate production and histone Kla by hypoxia was attenuated by a lactate dehydrogenase inhibitor (oxamate) or a PDK1 inhibitor (DCA) (Extended Data Fig. 3i, j). Deletion of both *LDHA* and *LDHB* fully suppressed production of lactate and histone Kla in HepG2 cells under normoxic conditions (Extended Data Fig. 3k, l). Owing to poor cell viability, hypoxic conditions could not be tested (data not shown).

Emerging evidence shows that lactate has regulatory functions in both innate and adaptive immune cells[8] and induces marked changes in gene expression[9], suggesting that lactate is not simply a 'waste product' of glycolysis. Pro-inflammatory M1 macrophages undergo metabolic reprogramming towards aerobic glycolysis, resulting in lactate production, whereas anti-inflammatory M2 macrophages trigger a metabolic program of increased oxidative phosphorylation and fatty acid oxidation[10]. Our discovery of histone Kla marks and their dynamics therefore suggests a role in regulating gene expression during M1 macrophage polarization.

To test this hypothesis, we examined the dynamics of lactate production and histone Kla marks during M1 macrophage polarization after treatment of BMDMs with lipopolysaccharide (LPS) and interferon-γ (IFNγ). We observed increased intracellular lactate levels 16–24 h after M1 activation (Fig. 3a), which correlated with increased histone Kla levels (Fig. 3b, c). By contrast, histone Kac levels were decreased at these time points (Fig. 3b, c). This differential pattern was confirmed by U-$^{13}C_6$-glucose labelling experiments, in which $^{13}C$-labelled histone Kac peaked 3 h after labelling and declined to a steady state, whereas histone Kla increased over the 24-h time course (Extended Data Fig. 4a–d). In addition, the LDHA-specific inhibitor GNE-140 reduced $^{13}C$ incorporation into histone Kla, but not Kac (Extended Data Fig. 4e, f). The increase of histone Kla during M1 polarization is intrinsic and not due to paracrine effects, because replenishing cells with fresh media every 4 h did not affect Kla levels (Extended Data Fig. 4g). Increases in lactate production and histone Kla are also specific to M1 macrophages because they were not observed in M2-polarized BMDMs (Fig. 3d and Extended Data Fig. 4h), which are more reliant on fatty acid oxidation[10].

## Histone Kla induces M2-like genes in M1 macrophages

Histone modifications have an important role in the regulation of gene expression[11]. To investigate histone Kla-associated genes 24 h after M1 polarization of macrophages, we performed RNA sequencing (RNA-seq) and paired chromatin immunoprecipitation followed by sequencing (ChIP–seq) using anti-H3K18la or anti-H3K18ac antibodies (the specificities of which were validated by dot blots) (Extended Data Fig. 3a–d),

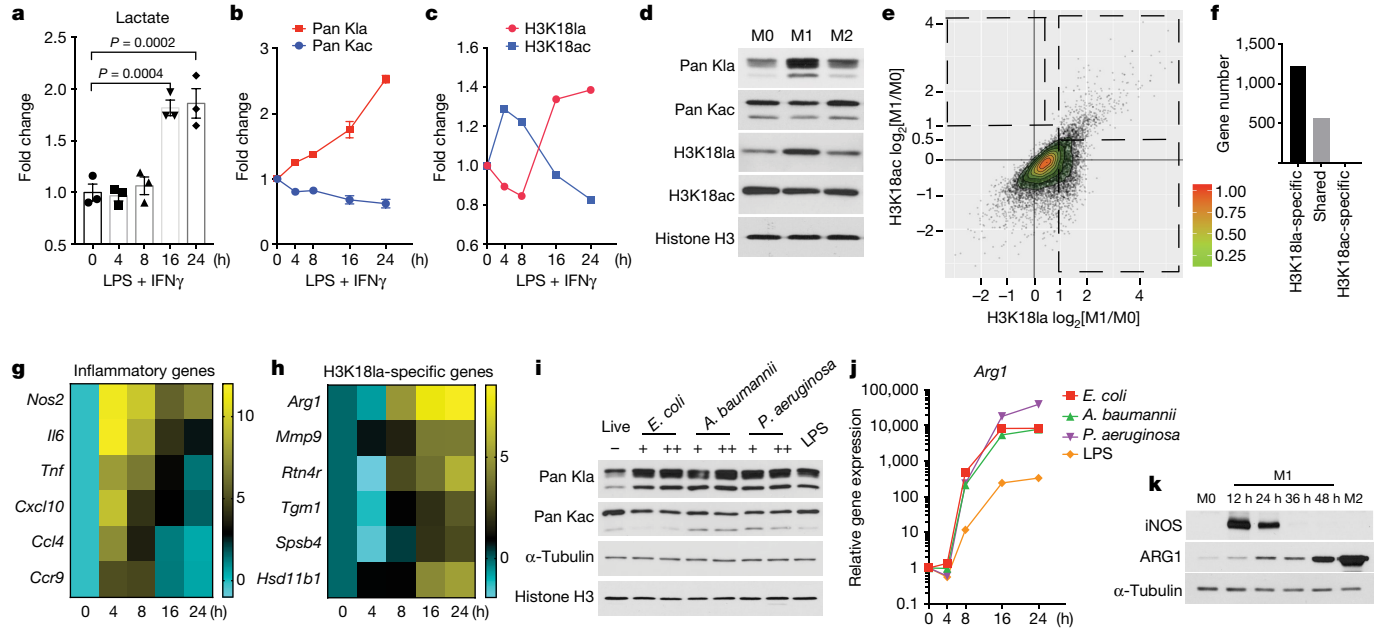



**Fig. 3 | Increased histone Kla during M1 macrophage polarization is associated with M2-like gene activation. a**–**c**, BMDMs were activated with LPS and IFNγ. **a**, Intracellular lactate levels were measured using a lactate colorimetric kit. $n = 3$ biological replicates; statistical significance was determined using one-way ANOVA followed by Dunnett's multiple comparisons test. **b**, **c**, Histone acylations were analysed by immunoblots using whole-cell lysates. ImageJ was used for quantification; $n = 3$ technical replicates. Data represent two independent experiments. **d**, BMDM cells were stimulated with PBS (M0), LPS and IFNγ (M1), and IL-4 (M2) for 24 h, respectively. Acid-extracted histones were used for immunoblots. **e, f**, Scatter plot (**e**) and bar plot (**f**) showing genes with promoters marked by exclusively increased H3K18la (H3K18la-$\log_2$[M1/M0] ≥ 1 and H3K18ac-$\log_2$[M1/M0] ≤ 0.5, H3K18la-specific); increased in both H3K18la and H3K18ac (H3K18la-$\log_2$[M1/M0] ≥ 1 and H3K18ac-$\log_2$[M1/M0] ≥ 0.5, shared); or exclusively increased H3K18ac (H3K18ac-$\log_2$[M1/M0] ≥ 1 and H3K18la-$\log_2$[M1/M0] ≤ 0.5, H3K18ac-specific). **g, h**, Heat maps showing gene expression kinetics (using reads per kilobase of transcript per million mapped reads (RPKM) values from RNA-seq) of exemplar inflammatory genes (**g**) and H3K18la-specific genes (**h**). The colour key represents $\log_2$-transformed fold change relative to gene expression at 0 h. $n = 4$ biological replicates. **i, j**, BMDM cells were infected with indicated Gram-negative bacteria or LPS, respectively. **i**, Histone Kla levels were measured by immunoblot at 24 h after bacterial challenge. '+' indicates lower dose, and '++' indicates higher dose. **j**, Gene expression was analysed by quantitative PCR with reverse transcription (RT–qPCR) at indicated time points after bacterial challenge. $n = 3$ biological replicates. **k**, Protein levels of inducible nitric oxide synthase (iNOS) and ARG1 were analysed by immunoblots from BMDMs activated by the indicated stimuli. Data in **a**–**c**, **j** are mean and s.e.m. Data in **d**, **i** and **k** represent three independent experiments.

ChIP and quantitative PCR (qPCR) assays (Extended Data Fig. 4i, j) and immunoblots (Extended Data Fig. 4k).

Our ChIP–seq data showed that H3K18la and H3K18ac were both enriched in promoter regions (±2 kb around transcriptional start sites) (Extended Data Fig. 4l) and were indicative of steady-state mRNA levels (Extended Data Fig. 4m, n). In addition, increased H3K18la (twofold increase) marked more genes than decreased H3K18la (twofold decrease), whereas the converse was true for the H3K18ac modification (Fig. 3e). Moreover, most genes marked by increased H3K18la were specific, because 68% of these genes (1,223 out of 1,787) did not display significantly increased H3K18ac (Fig. 3e, f and Supplementary Tables 2, 3). By contrast, no H3K18ac-specific genes were identified (Fig. 3e, f). Representative tracks from ChIP–seq studies are shown in Extended Data Fig. 4o, p.

To study correlations between H3K18la marks and gene expression, we performed RNA-seq analysis 0, 4, 8, 16 and 24 h after challenge with LPS and IFNγ (Extended Data Fig. 5a and Supplementary Table 4). As expected, inflammatory response genes (for example, *Nos2*) were induced as early as 4 h after challenge with LPS and IFNγ, and their expression levels steadily declined at later time points (Fig. 3g). Notably, the 1,223 genes specifically marked by increased H3K18la were more likely to be activated or reactivated at later time points (16 or 24 h) during M1 polarization (Fig. 3h and Extended Data Fig. 5a–c), which correlated well with the induction of intracellular lactate and histone Kla levels at these later time points (Fig. 3a–c). Gene Ontology (GO) analysis revealed that these H3K18la-specific genes were enriched in biological pathways that are independent of inflammation (Extended Data Fig. 5d). One of these enriched pathways was wound healing (for example, *Arg1*), which has been associated with the M2-like phenotype (Fig. 3h and Extended Data Fig. 5d). To corroborate these findings with more physiologically relevant stimuli, we treated BMDMs (M0) with live or dead Gram-negative bacteria (*Escherichia coli*, *Acinetobacter baumannii* and *Pseudomonas aeruginosa*) to stimulate M1 polarization. Similar to treatment with LPS, bacteria induced lactate production and global histone Kla but not histone Kac levels (Fig. 3i and Extended Data Fig. 5e, f), and the kinetics of early cytokine and late *Arg1* expression were maintained (Fig. 3j and Extended Data Fig. 5g–j).

Arginine metabolism is a key catabolic and anabolic process that is regulated during macrophage polarization. M1 macrophages are thought to have low levels of ARG1 and metabolize arginine via nitric oxide synthase to produce nitric oxide to kill pathogens, whereas M2 macrophages have high levels of ARG1, which produces ornithine to facilitate wound healing[12]. Consistent with their RNA dynamics, ARG1 protein levels and activity were markedly increased 24–48 h after M1 polarization, whereas NOS2 protein levels and function peaked 12 h after M1 polarization and declined at later time points (Fig. 3k and Extended Data Fig. 5k). Collectively, these findings suggest that induction of lactate during M1 activation might promote a late-phase switch to a more homeostatic phenotype, which shares some similarity with the M2-like phenotype. Indeed, previous studies showed that treating BMDMs with lactate derived from tumour cells drives an M2-like phenotype that is characteristic of tumour-associated macrophages (TAMs)[13]. Using mouse cancer models, we observed a positive correlation between *Arg1* expression and histone Kla levels, but not histone Kac levels in TAMs

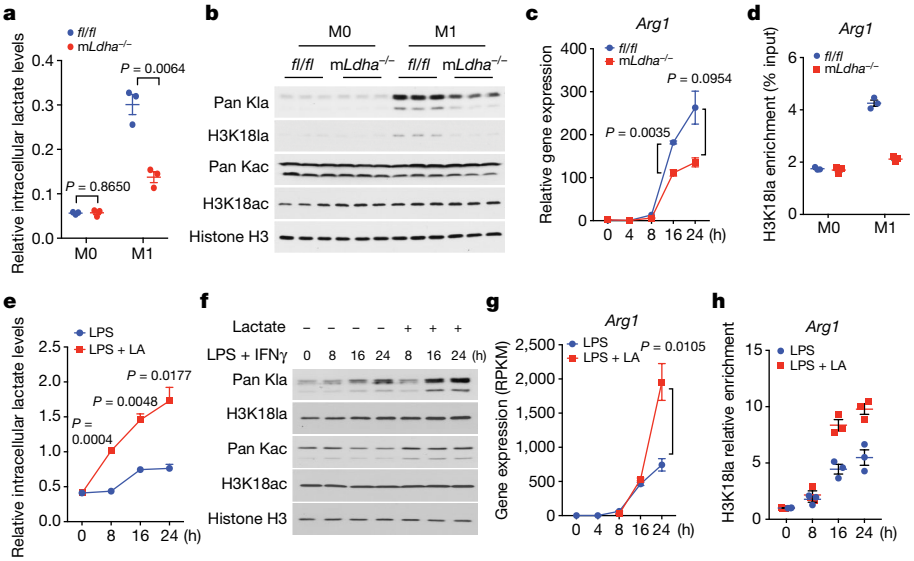



**Fig. 4 | Lactate activates M2-like gene expression through histone Kla.** **a**–**d**, Decreased lactate production in LDHA-deficient (myeloid-specific *Ldha*$^{-/-}$; m*Ldha*$^{-/-}$) BMDM cells resulted in lowered histone Kla levels and *Arg1* expression during M1 polarization. *fl/fl*, littermate control mice. Intracellular lactate levels were measured using a lactate colorimetric kit (**a**) and global histone Kla levels were measured by immunoblots (**b**) 24 h after M1 polarization. **c**, Gene expression was analysed by RT–qPCR at indicated time points after M1 polarization. **a**–**c**, $n = 3$ biological replicates. **d**, H3K18la occupancy was analysed by ChIP–qPCR 24 h after M1 polarization. Data represent three technical replicates from pooled samples. **e**–**h**, Exogeneous lactic acid (LA) (25 mM) was added to BMDM cells 4 h after M1 polarization (LPS + IFNγ), and cells were collected at indicated time points after M1 polarization for intracellular lactate measurement (**e**), histone Kla immunoblot analysis (**f**), gene expression analysis (**g**) and H3K18la occupancy analysis by ChIP–qPCR (**h**). **e**, $n = 3$ biological replicates. **f**, Data represent three independent experiments. **g**, $n = 4$ biological replicates. **h**, Data represent three technical replicates from pooled samples. Data in **a**, **c**–**e**, **g** and **h** are mean and s.e.m.; statistical significance was determined using multiple *t*-tests corrected using the Holm–Sidak method (**a**, **c**, **e**, **g**).

isolated from B16F10 melanoma and LLC1 lung tumours (Extended Data Fig. 6a–e).

Changes in gene expression during M1 polarization are caused by complex signalling cascades induced by LPS and IFNγ, including the induction of lactate and histone Kla. To substantiate the role of lactate and histone Kla in the regulation of gene expression, we manipulated levels of lactate during M1 polarization and examined its effect on the expression of *Arg1*, an M2-like gene. We first lowered lactate levels by deleting *Ldha* (LysM-Cre$^{+/-}$*Ldha*$^{fl/fl}$; Extended Data Fig. 7a–c). Lactate production and global histone Kla levels were both decreased in LDHA-deficient macrophages during M1 polarization (Fig. 4a, b). Although deleting *Ldha* in macrophages did not alter the expression of pro-inflammatory cytokines (Extended Data Fig. 7d–g), it attenuated *Arg1* and decreased histone Kla marks at the *Arg1* promoter (Fig. 4c, d). Similar findings were obtained when macrophages were M1 polarized in the presence of glycolysis inhibitors (2-DG, DCA and GNE-140) (Extended Data Fig. 7h–m). Next, we increased lactate levels by treating M1 macrophages with exogenous lactate. Exogenous lactate increased intracellular lactate (Fig. 4e) and histone Kla levels (Fig. 4f), and induced *Arg1* expression (Fig. 4g) and Kla levels at the *Arg1* promoter (Fig. 4h). By contrast, exogenous lactate did not affect the expression of early pro-inflammatory genes (Extended Data Fig. 8a–d). In addition, exogenous lactate enhanced the expression of other M2-like genes, such as *Vegfa* during M1 polarization (Extended Data Fig. 8e–h and Supplementary Table 5). Thus, these data confirmed the positive role of lactate and histone Kla in driving expression of M2-like genes during M1 macrophage polarization.

## Histone Kla directly stimulates gene transcription

Our observed correlations between lactate, H3K18la and M2-like gene expression does not necessarily indicate that the histone Kla mark was a causative factor. Previous studies showed that exogenous lactate can alter *Arg1* and *Vegfa* expression in unstimulated (M0) macrophages by HIF1α[13]. However, HIF1α is unlikely to be important for regulating *Arg1* and *Vegfa* during M1 polarization as HIF1α protein was induced at early time points and bound to promoters of glycolytic genes but not *Arg1* and *Vegfa* (Extended Data Fig. 8i–m).

To examine whether histone Kla has a direct role in transcriptional regulation, we took advantage of a cell-free, recombinant chromatin-templated histone modification and transcription assay (Extended Data Fig. 9a) that was used previously to demonstrate direct transcriptional activation by p53- and p300-dependent histone Kac[14]. This assay, in which acetyl-CoA was replaced by L-lactyl-CoA (validated by HPLC and mass spectrometry; Extended Data Fig. 9h–k), demonstrated robust p53-dependent, p300-mediated H3 and H4 lactylation (Extended Data Fig. 9b) and a corresponding effect on transcription (Extended Data Fig. 9c). The effects paralleled those observed for acetyl-CoA-dependent histone acetylation and transcription. To confirm that transcription was directly mediated by lactylation of histones, rather than other proteins in the nuclear extract, recombinant chromatin was reconstituted with core histones bearing lysine-to-arginine mutations in histone tails[15]. Compared with wild-type histones, the H3 and H4 mutations, but not the H2A or H2B mutations, eliminated p300- and p53-dependent transcription (Extended Data Fig. 9d). Together, these findings suggest that, similar to histone acetylation, histone lactylation can directly promote gene transcription under the described conditions. To examine the potential activity of p300 as a histone Kla writer in cells, we overexpressed p300 in HEK293T cells and observed a modest increase in histone Kla levels (Extended Data Fig. 9e). By contrast, p300 deletion in HCT116 and HEK293T cells decreased histone Kla levels (Extended Data Fig. 9f, g). Although we cannot exclude an indirect effect by p300 in these cells, together with the in vitro enzymatic results, these data suggest that p300 is a potential histone Kla writer protein.

In response to bacterial infection, macrophages must react rapidly with a substantial pro-inflammatory burst to help kill bacteria and recruit additional immune cells to the infection site. During this process,

macrophages switch to aerobic glycolysis[10], which is thought to support pro-inflammatory cytokine expression during M1 activation[16] and produce the Warburg effect. Over time, this metabolic switch also increases intracellular lactate, which we show stimulates histone lysine lactylation 16–24 h after exposure to M1-polarizing stimuli. Histone lactylation is not required for the induction or suppression of pro-inflammatory genes. Instead, it serves as a mechanism to initiate expression of homeostatic genes that have been traditionally associated with M2-like macrophages. Our studies support a model in which the switch to aerobic glycolysis that occurs during M1 polarization starts a 'lactate timer' that uses an epigenetic mechanism to induce M2-like characteristics in the late phase, perhaps to assist with repairing collateral damage incurred by the host during infection.

High levels of lactate (for example, 40 mM in certain type of tumour tissue[17]) is also associated with major hallmarks of diseases such as cancer. Given that the Kla modification can be stimulated by lactate and contribute to gene expression, the Kla modification is likely to fill an important knowledge gap in our understanding of diverse physiopathology (for example, infection, cancer) with which lactate is intimately associated.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1678-1.


1. Pavlova, N. N. & Thompson, C. B. The emerging hallmarks of cancer metabolism. *Cell Metab.* **23**, 27–47 (2016).
2. Palsson-McDermott, E. M. & O'Neill, L. A. The Warburg effect then and now: from cancer to inflammatory diseases. *BioEssays* **35**, 965–973 (2013).
3. Sabari, B. R., Zhang, D., Allis, C. D. & Zhao, Y. Metabolic regulation of gene expression through histone acylations. *Nat. Rev. Mol. Cell Biol.* **18**, 90–101 (2017).
4. Kaelin, W. G., Jr & McKnight, S. L. Influence of metabolism on epigenetics and disease. *Cell* **153**, 56–69 (2013).
5. Tan, M. et al. Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* **146**, 1016–1028 (2011).
6. Liu, X. et al. Acetate production from glucose and coupling to mitochondrial metabolism in mammals. *Cell* **175**, 502–513 (2018).
7. Semenza, G. L. Oxygen sensing, hypoxia-inducible factors, and disease pathophysiology. *Annu. Rev. Pathol.* **9**, 47–71 (2014).
8. Haas, R. et al. Intermediates of metabolism: from bystanders to signalling molecules. *Trends Biochem. Sci.* **41**, 460–471 (2016).
9. Martinez-Outschoorn, U. E. et al. Ketones and lactate increase cancer cell "stemness," driving recurrence, metastasis and poor clinical outcome in breast cancer: achieving personalized medicine via Metabolo-Genomics. *Cell Cycle* **10**, 1271–1286 (2011).
10. Galván-Peña, S. & O'Neill, L. A. Metabolic reprograming in macrophage polarization. *Front. Immunol.* **5**, 420 (2014).
11. Allis, C. D. & Jenuwein, T. The molecular hallmarks of epigenetic control. *Nat. Rev. Genet.* **17**, 487–500 (2016).
12. Rath, M., Müller, I., Kropf, P., Closs, E. I. & Munder, M. Metabolism via arginase or nitric oxide synthase: two competing arginine pathways in macrophages. *Front. Immunol.* **5**, 532 (2014).
13. Colegio, O. R. et al. Functional polarization of tumour-associated macrophages by tumour-derived lactic acid. *Nature* **513**, 559–563 (2014).
14. An, W., Kim, J. & Roeder, R. G. Ordered cooperative functions of PRMT1, p300, and CARM1 in transcriptional activation by p53. *Cell* **117**, 735–748 (2004).
15. Tang, Z. et al. SET1 and p300 act synergistically, through coupled histone modifications, in transcriptional activation by p53. *Cell* **154**, 297–310 (2013).
16. Tannahill, G. M. et al. Succinate is an inflammatory signal that induces IL-1β through HIF-1α. *Nature* **496**, 238–242 (2013).
17. Walenta, S. et al. High lactate levels predict likelihood of metastases, tumor recurrence, and restricted patient survival in human cervical cancers. *Cancer Res.* **60**, 916–921 (2000).

# Methods

### Materials

Pan anti-Kac (PTM-101), pan anti-Kla (PTM-1401), anti-H3K18la (PTM-1406), anti-H4K5la (PTM-1407) and anti-H4K8la (PTM-1405) antibodies were generated by PTM Bio Inc.; anti-histone H3 (ab12079), anti-H3K18ac (ab1191) and anti-H3K27ac (ab4729) antibodies were purchased from Abcam; *Drosophila* spike-in antibody (61686) and spike-in chromatin (53083) were obtained from Active Motif; anti-LDHA (2012S) antibody was from Cell Signaling Technology; anti-α-tubulin (05-829) and anti-LDHB (ABC927) antibodies were from Millipore Sigma; anti-HIF-1a (NB100-105) antibody was from Novus Biologicals; anti-iNOS (GTX130246) and anti-ARG1 (GTX109242) antibodies were purchased from GeneTex; anti-p300 (sc-584) was from Santa Cruz Biotechnology; anti-CD11b monoclonal antibody (M1/70), PE-cyanine7 (25-0112-82) and anti-F4/80 monoclonal antibody (BM8), APC (17-4801-82) were from Thermo Fisher Scientific; lipopolysaccharides from *Escherichia coli* O111:B4 (L4391), sodium L-lactate (71718), L-(+)-lactic acid (L6402), sodium dichloroacetate (347795), cobalt (II) chloride hexahydrate (C8661), rotenone (R8875), and acetyl-CoA (A2056) were purchased from Sigma-Aldrich; sodium L-lactate (13C3, 98%) (CLM-1579-PK) and D-glucose (U-13C6, 99%) (CLM-1396-1) were purchased from Cambridge Isotope Laboratories. Recombinant mouse IFNγ protein (485-MI-100) was from R&D Systems; mouse IL-4 (130-097-760) was from Miltenyi Biotec; modified sequencing-grade trypsin was from Promega; lactate colorimetric assay kit II (K627-100), arginase activity colorimetric assay kit (K755-100), and nitric oxide synthase (NOS) activity assay kit (K205-100) were purchased from Biovision.

### Cell culture

MCF-7, MDA-MB-231, HeLa, A549, HepG2, MEF and RAW 264.7 cells were obtained from the American Type Culture Collection and cultured in DMEM supplemented with 10% FBS and 1% GlutaMAX (GIBCO). Cells were routinely tested for mycoplasma contamination (MP0035, Sigma-Aldrich), and only negative cells were used in experiments. No specific cell line authentication was performed. For growth under hypoxic conditions, cells were grown in a specialized, humidified chamber equilibrated with 1% oxygen, 94% nitrogen, 5% carbon dioxide for the indicated time.

### Mouse experiments

All animal use and experiments performed were approved by Institutional Animal Care and Use Committee (ACUP 72209) at the University of Chicago. *Ldha*$^{fl/fl}$ mice (Jackson Laboratory, 030112) and LysM-Cre mice (Jackson Laboratory, 004781) were used to generate LysM-Cre$^{+/-}$ *Ldha*$^{fl/fl}$ and littermate control LysM-Cre$^{-/-}$ *Ldha*$^{fl/fl}$ mice. The following primers were used for genotyping: *Ldha* forward: CTGAGCACACCCATG TGAGA and *Ldha* reverse: AGCAACACTCCAAGTCAGGA. *LysM-cre* (*LysM* is also known as *Lyz2*): CCCAGAAATGCCAGATTACG, *LysM* common: CTTGGGCTGCCAGAATTTCTC and *LysM* WT: TTACAGTCGGCCAGGC TGAC. Macrophages were derived from bone marrow of 8-week-old male C57BL/6 mice following the published procedure[18]. To induce an M1 or M2 phenotype, BMDM cells were stimulated with 5 ng ml$^{-1}$ of LPS and 12 ng ml$^{-1}$ of IFNγ, or 20 ng ml$^{-1}$ of IL-4, for 24 h or the indicated time. To infect BMDM cells with bacteria, overnight cultures of *E. coli*, *A. baumannii* or *P. aeruginosa* were diluted in RPMI-1640 and added to BMDM cells in 6-well plates at 2 and 20 multiplicity of infection. A control plate was either infected with paraformaldehyde-killed bacteria or treated with 5 ng ml$^{-1}$ LPS in the absence of bacteria. The plates were centrifuged at 975*g* for 30 min to promote infection, followed by a 30 min incubation in a humidified incubator at 37 °C at 5% $CO_2$. To kill extracellular bacteria, the medium overlying the confluent cell monolayer was replaced with fresh medium containing gentamicin at 100 μg ml$^{-1}$ and the plates were further incubated for 1 h. After incubation, media were removed from infected cells and replaced with fresh media containing 25 μg ml$^{-1}$ of gentamicin. For consistency, LPS-treated cells and cells infected with dead bacteria were also treated with gentamicin. Cells were cultured for 24h before lysis. Allocation of BMDM cells into different treated groups was randomized and not blinded.

### Tumour inoculation and TAM isolation

LLC1 cells ($0.5 \times 10^6$) or B16F10 cells ($1 \times 10^6$) were injected into 7-week-old C57BL/6 mice (Jackson Laboratory). Once tumours reached approximately 600 mm$^3$, mice were killed for tumour isolation. Tumours were digested with type 4 collagenase (Worthington, 3 mg ml$^{-1}$) and hyaluronidases (Sigma, 1.5 mg ml$^{-1}$) in 1% BSA/PBS at 37 °C with shaking at 200 r.p.m. for 30 min. The digested tumour was then filtered through a 70-μm cell strainer, followed by red blood cell lysis step and passing through another 40-μm strainer. Cells were resuspended into isolation buffer (0.1% BSA/PBS, 2 mM EDTA), layered onto Ficoll-Paque PLUS (GE Healthcare), and centrifuged at 450*g* for 30 min without a break. Mononuclear immune cells were obtained by taking out the middle white layer. TAMs were then isolated using CD11b Microbeads (Mitenyi Biotec) as the company instructed. The purity of TAMs purity was confirmed by flow cytometry using CD11b and F4/80 antibody. Data were quantified by FlowJo v.10.4.1.

### Peptide immunoprecipitation

Histones from human MCF-7 or mouse BMDM cells were extracted using a standard acid-extraction protocol[19], and subjected to trypsin digestion as per the manufacturer's instructions. Pan anti-Kla or pan anti-Kac antibodies were first conjugated to Protein A Sepharose beads (GE Healthcare BioSciences) and then incubated with tryptically digested histone peptides with gentle agitation overnight at 4 °C. The beads were then washed three times with NETN buffer (50 mM Tris-Cl pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5% NP-40), twice with ETN buffer (50 mM Tris-Cl pH 8.0, 100 mM NaCl, 1 mM EDTA) and once with water. Peptides were eluted from the beads with 0.1% TFA and dried in a SpeedVac system (Thermo Fisher Scientific).

### HPLC–MS/MS analysis

The peptide samples were loaded onto a homemade capillary column (10 cm length × 75 mm ID, 3 μm particle size, Dr. Maisch GmbH) connected to an EASY-nLC 1000 system (Thermo Fisher Scientific). Peptides were separated and eluted with a gradient of 2% to 90% HPLC buffer B (0.1% formic acid in acetonitrile, v/v) in buffer A (0.1% formic acid in water, v/v) at a flow rate of 200 nl min$^{-1}$ over 60 min (34 min for coelution studies). The eluted peptides were then ionized and analysed by a Q-Exactive mass spectrometer (Thermo Fisher Scientific). Full mass spectrometry was acquired in the Orbitrap mass analyser over the range $m/z$ 300 to 1,400 with a resolution of 70,000 at $m/z$ 200. The 12 most intense ions with charge ≥2 were fragmented with normalized collision energy of 27 and tandem mass spectra were acquired with a mass resolution of 17,500 at $m/z$ 200.

### Isotopic-labelling experiments

MCF-7 cells were cultured in DMEM high-glucose media plus 10% FBS. To be labelled by isotopic lactate, cells were treated with 10 mM of $^{13}C_3$ sodium L-lactate for 24 h. To be labelled by isotopic glucose, cells were switched to DMEM No-Glucose media (Gibco) for 24 h, followed by supplementation with 25 mM of U-$^{13}C_6$ D-glucose and continued culturing for three passages. Histones were extracted, digested with trypsin, immunoprecipitated using a pan anti-Kla antibody and analysed by HPLC–MS/MS as described above.

### SILAC-based quantification

MCF-7 cells were cultured in either 'heavy' (L-Lys-$^{13}C_6$, $^{15}N_2$) or 'light' (L-Lys-$^{12}C_6$, $^{14}N_2$) DMEM, supplemented with 10% dialysed FBS (Serum Source International Inc.), for more than six passages, to achieve more than 99% labelling efficiency. Heavy-labelled and light-labelled cells were mixed in a 1:1 ratio. Histones were extracted, digested with trypsin,

immunoprecipitated using a pan anti-Kla antibody, and analysed by HPLC–MS/MS as described above. Quantification was analysed by Maxquant[20]. Ratio H/L derived from Maxquant was then normalized by protein abundance.

### Synthesis of L-lactyl-CoA

L-Lactic acid (90 mg, 1 mmol) was dissolved in 5 ml of freshly distilled $CH_2Cl_2$. *N*-hydroxysuccinimide (115 mg, 1 mmol) was added to this solution, and the reaction mixture was sonicated to obtain a clear solution. Then, *N*,*N*′-dicyclohexylcarbodiimide (DCC, 227 mg, 1.1 mmol) was added. A white precipitate formed after addition. The reaction mixture was stirred at room temperature overnight. Then the white precipitate was filtered and washed with $CH_3CN$. The resulting organic solvent was evaporated by vacuum to afford crude product L-lactyl-NHS (170 mg, 91% yield), which was used in the next step without further purification. CoA hydrate (0.0065 mmol; 5 mg) was dissolved in 1.5 ml of 0.5 M $NaHCO_3$ (pH 8.0) and cooled down on ice bath. Then, L-lactyl-NHS (2.5 mg, 0.013 mmol) in 0.5 ml of $CH_3CN$/acetone (1:1 v/v) was added dropwise to the CoA solution. The reaction solution was stirred at 4 °C overnight and then quenched by adjusting pH to 4.0 with 1.0 M HCl. The reaction mixture was then subjected to RP-HPLC purification with gradient 5–45% buffer A in buffer B over 30 min at flow rate 5 ml $min^{-1}$; UV detection wavelength was fixed at 214 and 254 nm (HPLC buffer A: 0.05% TFA in water; HPLC buffer B: 0.05% TFA in acetonitrile). The fractions were collected and lyophilized after flash-freeze with liquid nitrogen. m = 2 mg, yield 38% $^1$H NMR (400 MHz, Deuterium oxide) δ 8.57 (s, 1H), 8.33 (s, 1H), 6.12 (d, *J* = 5.7 Hz, 1H), 4.49 (s, 1H), 4.29 – 4.24 (m, 1H), 4.14 (s, 2H), 3.93 (s, 1H), 3.75 (d, *J* = 8.6 Hz, 1H), 3.48 (d, *J* = 7.6 Hz, 1H), 3.35 (t, *J* = 6.4 Hz, 2H), 3.22 (d, *J* = 5.2 Hz, 3H), 2.89 (q, *J* = 6.2 Hz, 2H), 2.32 (t, *J* = 6.4 Hz, 2H), 1.23 (d, *J* = 6.9 Hz, 3H), 0.83 (s, 3H), 0.70 (s, 3H). MALDI *m*/*z* calculated for $C_{24}H_{41}N_7O_{18}P_3S^+$ $[M+H]^+$: 840.1, found 839.6.

### In vitro chromatin template-based histone modification and transcription assays

Purification of recombinant proteins and chromatin assembly were performed as previously described[15]. The chromatin-templated histone modification and transcription assays were as described previously[15], except that lactyl-CoA was used in place of acetyl-CoA and [α-$^{32}$P]CTP was used in place of [α-$^{32}$P]UTP. The H3KR, H4KR, H2AKR and H2BKR histone mutants were the same as previously described[15]. Histone modifications were monitored by immunoblot and transcription products were monitored by autoradiography as described[15].

### RNA-seq

Total RNA was extracted from BMDM cells activated as indicated using a RNeasy Plus Mini Kit (74134, Qiagen). Two to four micrograms of total RNA were used as starting material to prepare libraries using Illumina TruSeq Stranded mRNA Library Prep Kit Set A (RS-122-2101, Illumina). The size of the libraries was selected by using the Agencourt AMPure XP beads (A63882, Beckman Coulter), with average size of 400 bp. The libraries were sequenced using Illumina HiSeq 4000 (pair end 50 bp).

Bioinformatic analysis of RNA-seq data: sequencing quality was evaluated by FastQC v.0.11.4. All reads were mapped to the reference genome of Illumina iGenomes UCSC mm10 using HISAT2 v.2.1.0[21]. Differential expression analysis was implemented using edgeR v.3.16.5[22], after retaining only genes for which counts per million (cpm) was larger than one in four samples and normalizing the library sizes across samples using the TMM method of the edgeR package. Hierarchical clustering was performed and heat maps were generated using Perseus v.1.6.1.1 (http://www.coxdocs.org/doku.php?id=perseus:start). The $\log_2$-transformed gene expression values (RPKM) were normalized by subtracting the mean in every row, and hierarchically clustered with a Pearson correlation algorithm. Gene Ontology analysis (GOTERM_BP_DIRECT) was carried out using DAVID bioinformatics resources 6.8[23,24].

The following primers were used for RT-qPCR analysis: *Arg1*: CTCC AAGCCAAAGTCCTTAGAG, AGGAGCTGTCATTAGGGACATC; *Vegfa*: CCACGACAGAAGGAGAGCAGAAGTCC, CGTTACAGCAGCCTGCACAG CG; *Il6*: GTTCTCTGGGAAATCGTGGA, TTTCTGCAAGTGCATCATCG; *Il1b*: TTTGACAGTGATGAGAATGACC, CTCTTGTTGATGTGCTGCTG; *Ifnb1*: CAGCTCCAAGAAAGGACGAAC, GGCAGTGTAACTCTTCTGCAT; *Cxcl10*: CCAAGTGCTGCCGTCATTTTC, GGCTCGCAGGGATGATTTCAA; *Tnfa*: CCCTCACACTCAGATCATCTTCT, GCTACGACGTGGGCTACAG; and *Rn18s* (18S rRNA): GTAACCCGTTGAACCCCATT, CCATCCAATCGGTAGTAGCG.

### ChIP–seq

Native ChIP was carried out following the published protocol[25] with spike-in for normalization purpose. Spike-in was carried out according to vendor protocols (61686, Active Motif). In brief, 50 ng of Spike-in chromatin (53083, Active Motif) was added to 25 μg of BMDM chromatin to incubate with 2 μg Spike-in antibody (61686, Active Motif) together with 4 μg of anti-H3K18la or anti-H3K18ac antibodies. After 4 h of incubation at 4 °C, Protein A Sepharose (17-5280-01, GE Healthcare Life Sciences) was added and incubated for another 2 h, followed by sequential wash with buffer TSE I (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.0, 150 mM NaCl), TSE II (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.0, 500 mM NaCl), buffer III (0.25 M LiCl, 1% NP-40, 1% deoxycholate, 1 mM EDTA, 10 mM Tris-HCl pH 8.0), and TE buffer (1 mM EDTA, 10 mM Tris-HCl pH 8.0). Chromatin DNA was finally eluted with buffer containing 1% SDS and 0.1 M $NaHCO_3$. The eluates were digested with RNase A (12091021, Thermo Fisher Scientific) and proteinase K (AM2546, Thermo Fisher Scientific). DNA was recovered using the QIAquick PCR purification kit (28106, Qiagen) according to the manufacturer's instructions.

ChIP–seq libraries were constructed with an Accel-NGS 2S Plus DNA Library Kit (Swift Biosciences) according to the manufacturer's protocol. The libraries were then amplified and assessed for fragment size using TapeStation (Agilent) and quantified using a Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific). The indexed libraries were pooled and sequenced on a Hiseq4000 Sequencer (Illumina) using the 50-nucleotide single-read configuration.

Bioinformatics analysis of ChIP–seq data: sequencing quality was evaluated by FastQC v.0.11.4. All reads were mapped to the reference genome of Illumina iGenomes UCSC mm10 using Bowtie v.2.2.6[26,27], and only uniquely mapped reads were retained. Then SAMtools v.0.1.19[28] was used to convert files to bam format, sort, and remove PCR duplicates. Peaks were called using MACS v.2.2.1[29] under $q = 0.01$. To quantify and directly compare H3K18la or H3K18ac in different samples (M0 and M1 macrophages), the uniquely mapped H3K18la or H3K18ac reads in promoter regions (± 2 kb around transcriptional start sites) of each gene were counted by featureCounts v.1.5.0-p1[30], and then normalized by Spike-in ChIP read counts of the corresponding condition (M0 or M1 macrophages). The overlap genes in ChIP–seq and RNA–seq data were used for all subsequent analysis. Gene Ontology analysis (GOTERM_BP_DIRECT) was carried out using DAVID Bioinformatics Resources 6.8[23,24].

The following primers were used for qPCR analysis of gene promoter regions in human cells: *FOXO3* (previously known as *FOXO3A*) promoter: CAGTGAGTGTGTGCAGCTTG, AAAGCCTCCTGTTTGTG CTT; *FOXO3* downstream: TGCACACAGAAGCCAGAAG, GCTCCCCA CAGAGACGTAA; *LDHA* promoter: TAAGGGTGGGGGATACCTCT, CCCAAGAGAAAAATGCAAGC. The following primers were used for qPCR analysis of gene promoter regions in mouse cells: *Arg1*/*Arg1*-PTM: AAGCTGTGGCCTCAGAACAT, GGTAACCGCTGTGAAAGGAT; *Arg1*-HRE-1kb: CCCGAGTTTGACCCGAAGAA, CTTTACACAGGGACC GGACC; *Arg1*-HRE-2kb: TGTCTCTCCCAGTTTCCCCA, AGCAACTTGG CATCTGATGGA; *Vegfa*/*Vegfa*-PTM: CGAGCTAGCACTTCTCCCAG, AACTTCTGGGCTCTTCTCGC; *Vegfa*-HRE-1kb: GGCACCAAATTTGTGG CACT, CTGCCAGACTACACAGTGCA; *Vegfa*-HRE-2kb: ACCTGATCC TGATCCCTGCT, CAGCCTCTGTTATGCCACGA; *Vegfa*-HRE-3kb: GCAGAACCTAGGCTTCACGT, TTGAAAGGGCTGACATGGCT; *Eno1*:

AAGGTCATCAGCAAGGTCGT, CGTACTCCGAGTCTCACACG; *Glut1* (also known as *Slc2a1*): TAGATCCCCTCCCTCTTGCT, GAACACGTAGCCTGC TCACA; gene desert: CTGCCAGGGTTGTAGAGAGG, GCCAGATCATATT GGCTTGG.

### Statistical analysis

No statistical methods were used to predetermine sample size. The significance of differences in the experimental data were determined using GraphPad Prism 7.0 software. All data involving statistics are presented as mean ± s.e.m. For data presented without statistics, experiments were repeated at least three times to ensure reproducibility, unless otherwise stated. The experiments were not randomized, and investigators were not blinded to allocation during experiments and outcome assessment.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

The ChIP–seq and RNA–seq data have been made available at the Gene Expression Omnibus (GEO) repository under the accession number GSE115354. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[31] partner repository with the dataset identifier PXD014870. All other data are available from the authors upon reasonable request.


18. Kratz, M. et al. Metabolic dysfunction drives a mechanistically distinct proinflammatory phenotype in adipose tissue macrophages. *Cell Metab.* **20**, 614–625 (2014).
19. Shechter, D., Dormann, H. L., Allis, C. D. & Hake, S. B. Extraction, purification and analysis of histones. *Nat. Protocols* **2**, 1445–1457 (2007).
20. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
21. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).
22. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
23. Huang, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protocols* **4**, 44–57 (2009).
24. Huang, W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009).
25. Cuddapah, S. et al. Native chromatin preparation and Illumina/Solexa library construction. *Cold Spring Harb. Protoc.* **2009**, pdb prot5237 (2009).
26. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
27. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
28. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
29. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
30. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
31. Perez-Riverol, Y. et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **47** (D1), D442–D450 (2019).

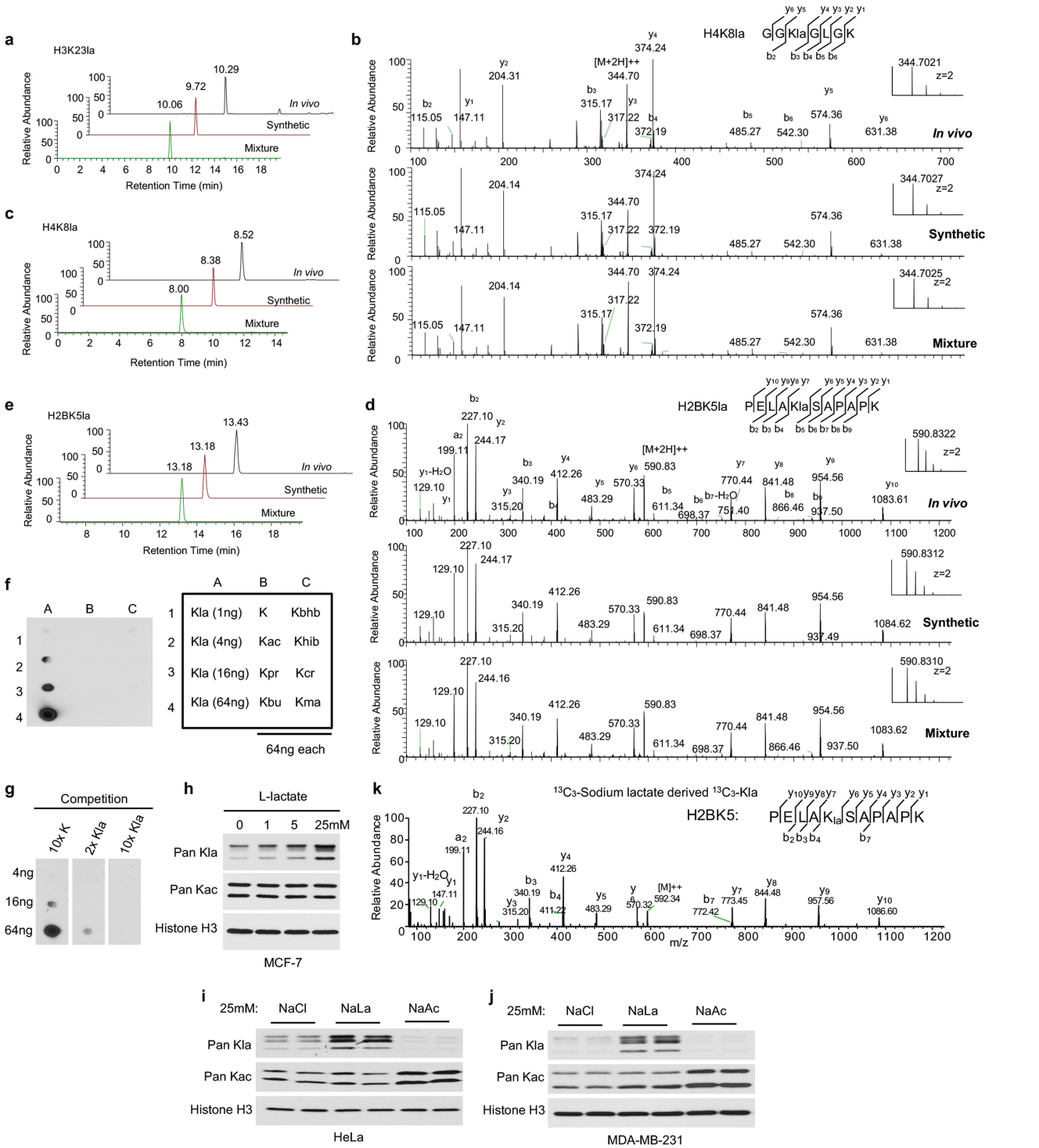


**Acknowledgements** HEK293T p300 knockout cells were provided by X. Li. We thank S. Khochbin for brainstorming and critical reading of this manuscript. We thank K. Delaney and all other members of the Zhao and Becker laboratories for discussions and technical support. This work was supported by the University of Chicago, Nancy and Leonard Florsheim family fund (Y.Z.), NIH grants R01GM115961, R01DK118266 (Y.Z.), R01DK102960, R01HL137998 (L.B.), R01CA129325, R01DK071900 (R.G.R.), and NSF1808087 (Y.G.Z.).

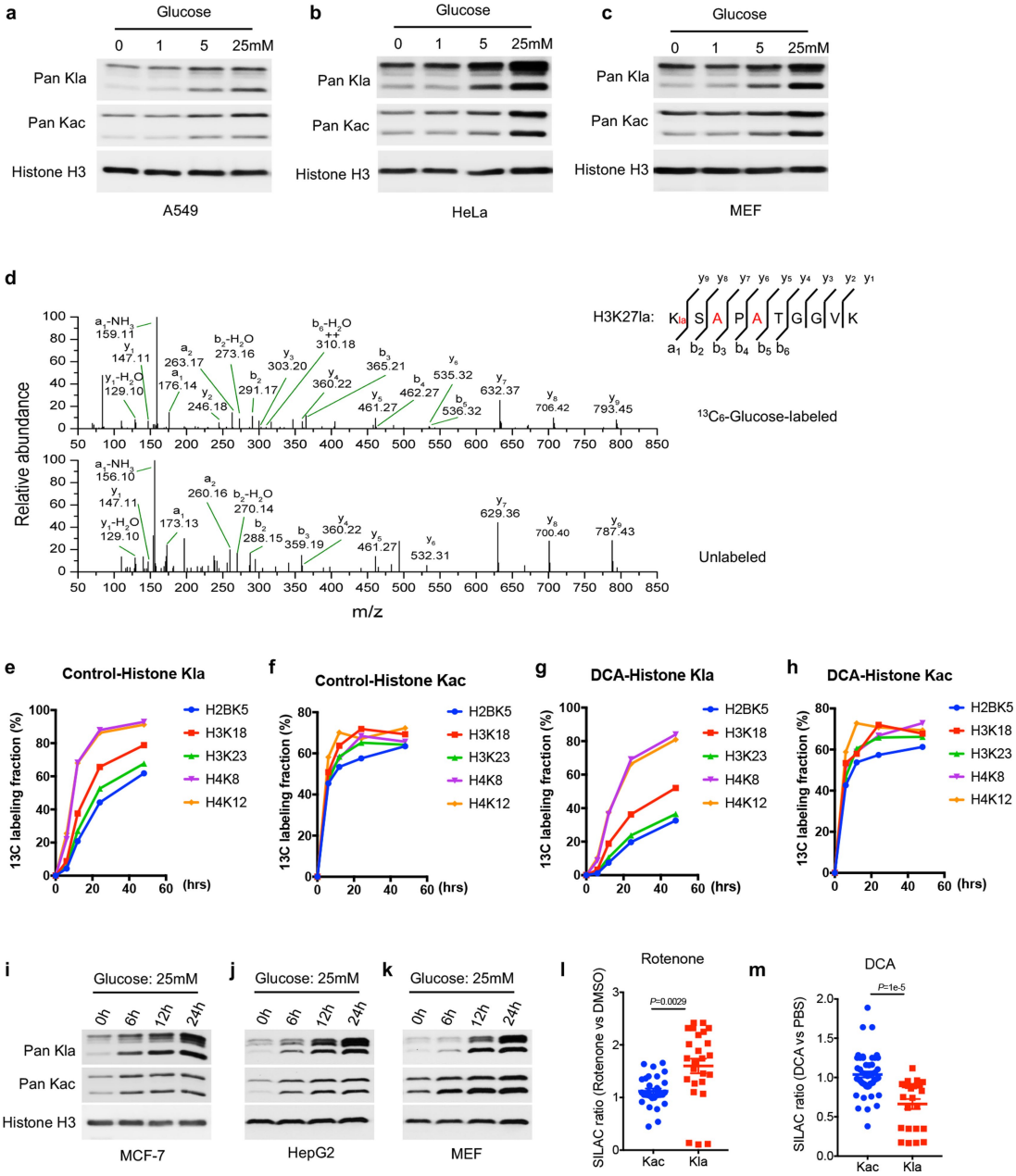

**Author contributions** Y.Z. conceived the project and developed the general ideas and research strategy. D.Z., L.B. and Y.Z. designed the experimental approach and composed the manuscript. D.Z. performed most of the experiments. Z.T. and R.G.R. carried out in vitro chromatin-based transcription experiments. Y.W., H.H., W.L., J.D., L.D., S.K., S.L. and M.P.-N. contributed to mass spectrometry-related experiments and analysis; R.H., Z.Y. and B.R. performed the library construction and next-generation sequencing for ChIP–seq and RNA-seq; M.H. and Y.G.Z. synthesized L-lactyl-CoA. H.H. and D.Z. analysed ChIP–seq and RNA-seq data. G.Z. provided all primary BMDM cell cultures. D.C. and H.A.S. carried out the bacterial infection experiments; C.C. carried out TAM experiments.

**Competing interests** Y.Z. is a co-founder, board member, and advisor to PTM Bio Inc. L.B. is a co-founder and CSO of rMark Bio Inc., and a founder and CEO of Onchilles Pharma Inc.

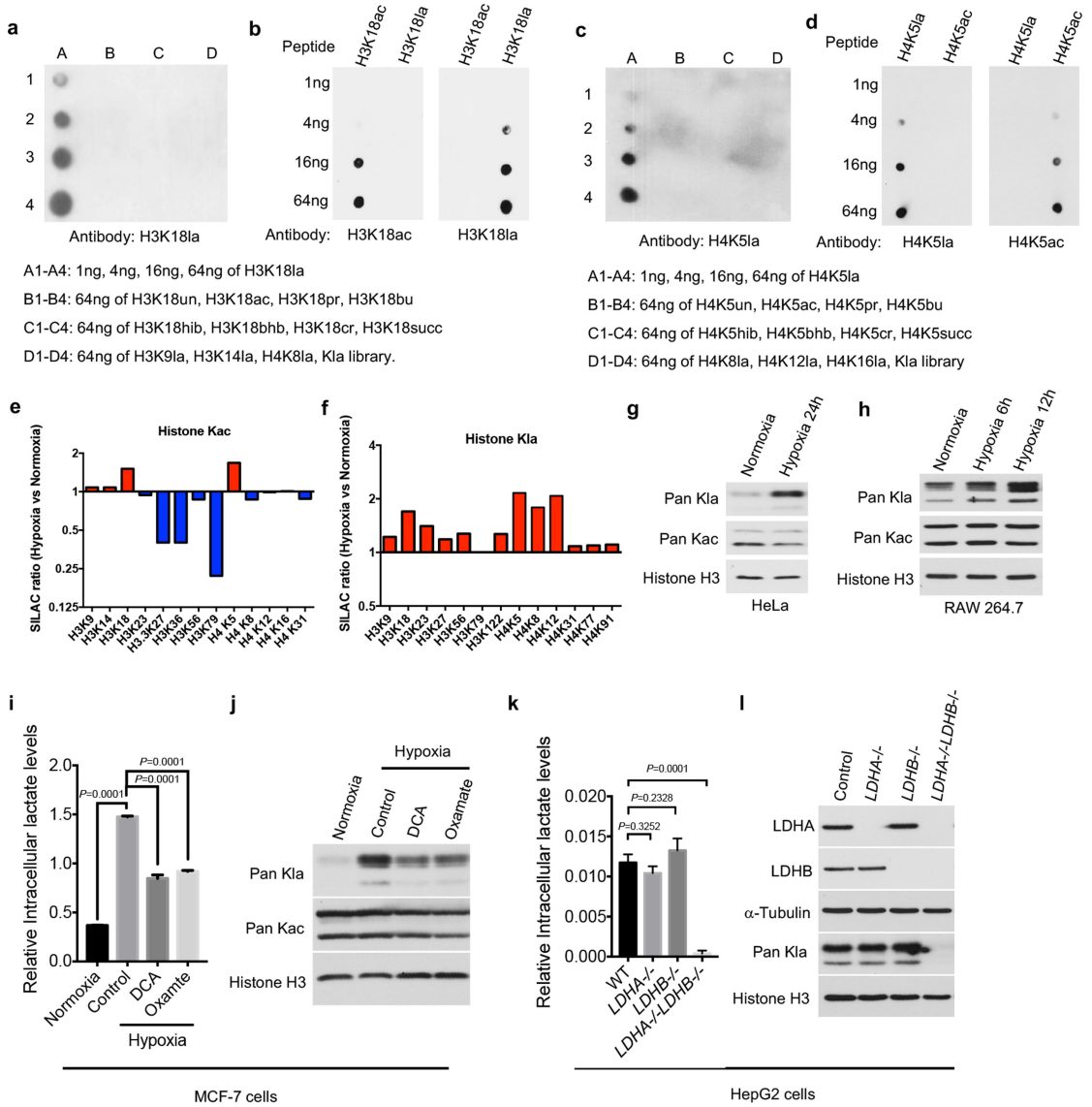

**Additional information**
**Supplementary information** is available for this paper at https://doi.org/10.1038/s41586-019-1678-1.
**Correspondence and requests for materials** should be addressed to L.B. or Y.Z.
**Peer review information** *Nature* thanks Luke O'Neill, Kathryn Wellen and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.
**Reprints and permissions information** is available at http://www.nature.com/reprints.



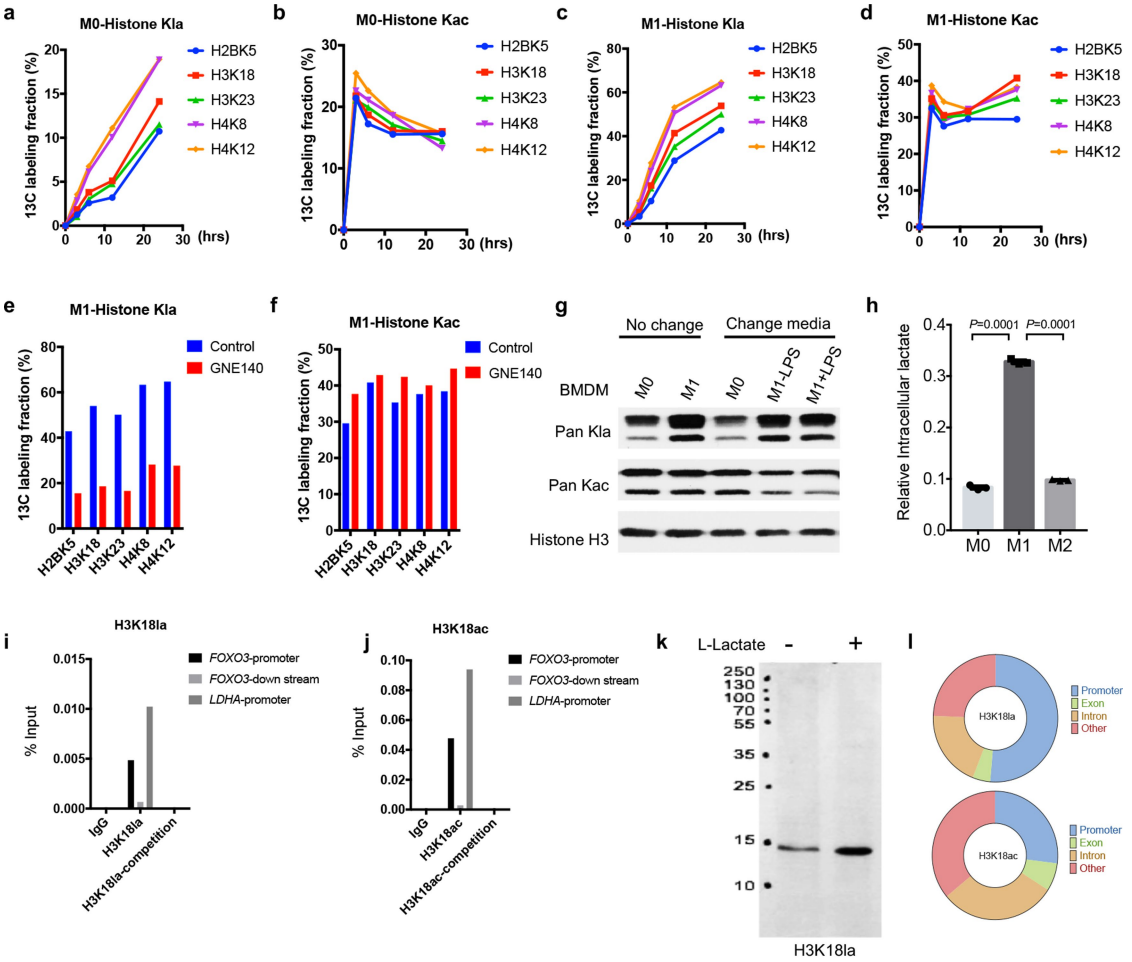

**Extended Data Fig. 1 | Validation of histone lysine lactylation.** **a**, **c**, **e**, Extracted ion chromatograms from HPLC–MS/MS analysis of histone Kla peptides derived from cultured cells (in vivo), the synthetic counterparts, and their mixtures. **b, d**, MS/MS spectra of histone Kla peptides derived from in vivo, the synthetic counterparts, and their mixtures. **f, g**, Antibody specificity tests by dot blot and competition assay. **f**, Dot blot was carried out with a pan anti-Kla antibody and the following peptide libraries. A1, A2, A3 and A4: dots contain 1, 4, 16 and 64 ng, respectively, of a peptide library containing a lactylated lysine residue. B1, B2, B3 and B4: dots contain 64 ng of a peptide library containing an unmodified (K), acetylated (Kac), propionylated (Kpr) and butyrylated (Kbu) lysine residue, respectively. C1, C2, C3 and C4: dots contain 64 ng of a peptide library containing a β-hydroxybutyrylated (Kbhb), 2-hydroxyisobutyrylated (Khib), crotonylated (Kcr) and malonylated (Kma) lysine residue, respectively. The libraries contained a mixture of CXXXKXXXX peptides, in which C is cysteine, X is a mixture of all 19 amino acids except for cysteine, and K is lysine with or without the indicated modifications. **g**, Competition was carried out by incubating the pan anti-Kla antibody with a twofold or tenfold excess of the indicated peptide libraries before the dot blot assay. **h**–**j**, Exogenous lactate boosts histone Kla levels. Immunoblot analysis of histone Kla and Kac from human MCF-7 cells treated with indicated doses of L-lactate (**h**), and from human HeLa (**i**) and MDA-MB-231 (**j**) cells treated with 25 mM sodium chloride, sodium lactate or sodium acetate. **k**, MS/MS spectra of an isotopically labelled histone Kla peptide identified from MCF-7 cells cultured with 10 mM isotopic ($^{13}C_3$) sodium L-lactate for 24 h. Data in **a**–**k** represent three independent experiments.




**Extended Data Fig. 2 | Histone Kla is modulated by the glycolysis pathway.** **a**–**c**, A549 (**a**), HeLa (**b**) and mouse embryonic fibroblast (MEF) (**c**) cells were cultured with indicated doses of glucose for 24 h, without pyruvate. Histone Kla and Kac were analysed by immunoblots using indicated antibodies. **d**, MS/MS spectra of a U-$^{13}C_6$-glucose labelled histone Kla peptide and its unlabelled counterpart from MCF-7 cells. **e**–**h**, Quantitative proteomic analysis of histone extracts from MCF-7 cells cultured in the presence of U-$^{13}C_6$ glucose for 6 h, 12 h, 24 h and 48 h, with or without 10 mM DCA. **i**–**k**, Histone Kla and Kac levels were analysed by immunoblots using whole-cell lysates from MCF-7, HepG2 and MEF cells exposed to 25 mM glucose for the indicated times. **l, m**, SILAC–MS/MS quantification of histone Kla and Kac marks from MCF-7 cells, comparing rotenone (10 nM, 24 h) versus DMSO treatment (**l**), and DCA (10 mM, 24 h) versus PBS treatment (**m**). SILAC ratio was normalized to protein abundance. Each dot in the scatter dot plot represents one identified peptide from core histone. Data are mean ± s.e.m. **l**, Kac: 1.121 ± 0.05084, $n = 31$; Kla: 1.599 ± 0.139, $n = 25$. **m**, Kac: 1.038 ± 0.03813, $n = 49$; Kla: 0.6627 ± 0.06376, $n = 24$. Statistical significance was determined using two-tailed Welch's $t$-test. Data in **a**–**d**, **i**–**k** represent three independent experiments. Data in **e**–**h** represent two independent experiments.

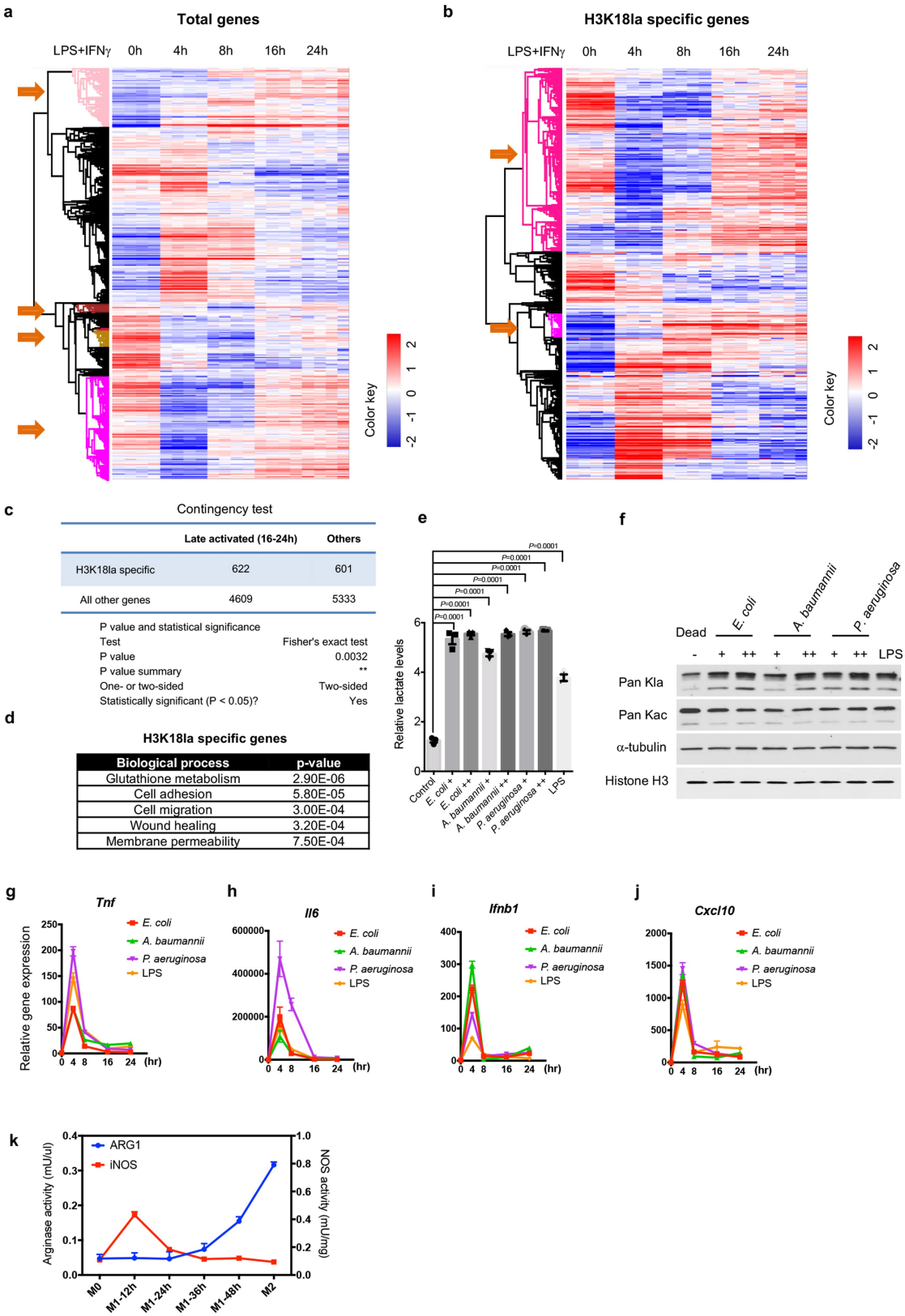



**Extended Data Fig. 3 | Histone Kla is induced by hypoxia. a–d**, Antibody specificity was analysed by dot blot assay. ac, acetyl lysine; bhb, β-hydroxybutyryl lysine; bu, butyryl lysine; cr, crotonyl lysine; hib, 2-hydroxyisobutyryl lysine; la, lactyl lysine; pr, propionyl lysine; succ, succinyl lysine; un, unmodified lysine. Kla library contains a mixture of CXXXKlaXXXX peptides, in which C is cysteine, X is a mixture of all 19 amino acids except for cysteine, and Kla is lactyl lysine. **e, f**, SILAC-MS/MS quantification of histone Kla and Kac marks from MCF-7 cells, comparing hypoxic (1% oxygen for 24 h) and normoxic conditions. SILAC ratio was normalized to protein abundance. **g, h**, Immunoblots of histone Kla and Kac from human HeLa and mouse RAW 264.7 cells in response to hypoxia (1% oxygen) at the indicated time. **i, j**, Intracellular lactate levels (**i**) and histone Kla levels (**j**) were measured in MCF-7 cells comparing normoxia, hypoxia (1% oxygen, 24 h) and hypoxia in the presence of 10 mM oxamate or DCA. **k, l**, Intracellular lactate levels (**k**) and histone Kla levels (**l**) were compared in *LDHA*$^{-/-}$, *LDHB*$^{-/-}$, *LDHA*$^{-/-}$*LDHB*$^{-/-}$ or wild-type (WT) HepG2 cells. Data are mean and s.e.m. from three biological independent samples; statistical significance was determined using one-way ANOVA followed by Dunnett's multiple comparisons test. Data in **a–d**, **g**, **h**, **k** and **l** represent three independent experiments.



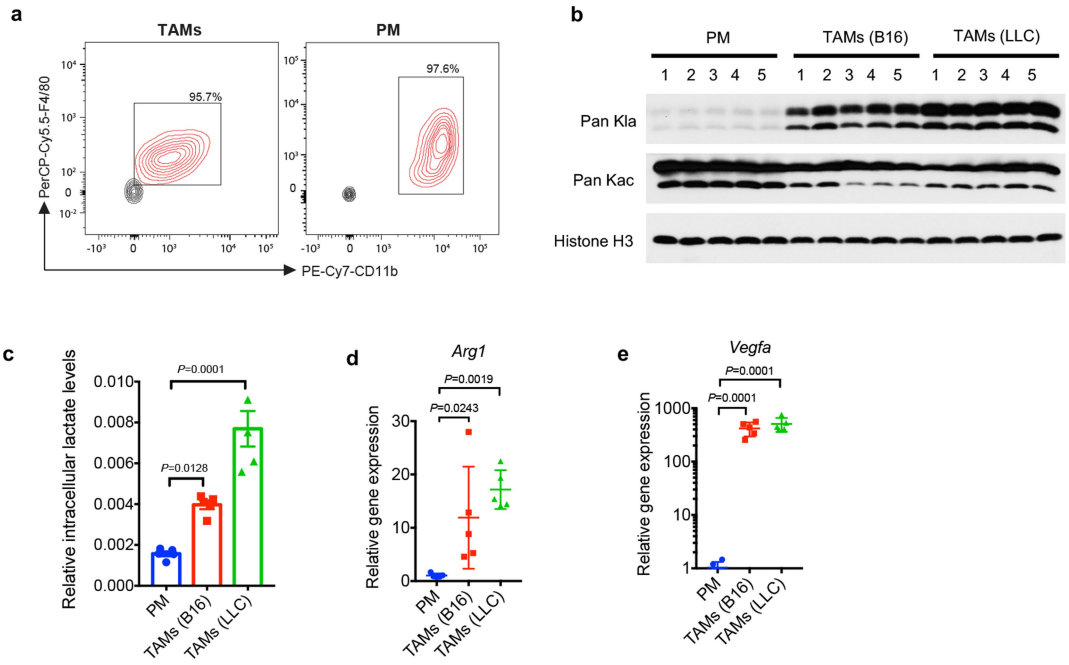

**Extended Data Fig. 4** | See next page for caption.

**Extended Data Fig. 4 | Histone Kla is induced during M1 macrophage polarization. a**–**f**, Quantitative proteomic analysis of histone extracts from M0 and M1 macrophages (BMDMs) cultured in the presence of U-$^{13}C_6$-glucose for 3, 6, 12 and 24 h, or with 10 μM GNE-140 (LDHA/B inhibitor) for 24 h. **g**, Histone Kla and Kac levels were analysed by immunoblots 24 h after activation by LPS and IFNγ, with or without replenishing fresh media (containing LPS and IFNγ or not) every 4 h. **h**, BMDM cells were stimulated with PBS (M0), LPS plus IFNγ (M1), and IL-4 (M2) for 24 h. Intracellular lactate was measured using a lactate colorimetric kit. Data are mean and s.e.m. from three biological independent samples; statistical significance was determined using one-way ANOVA followed by Dunnett's multiple comparisons test. **i, j**, Antibody specificity was evaluated by ChIP–qPCR. Competition was carried out by pre-incubating the indicated antibodies with a tenfold excess of corresponding peptides. **k**, H3K18la antibody specificity was shown by full immunoblot using total lysate from MCF-7 cells with or without 10 mM sodium L-lactate treatment for 24 h. **l**, H3K18la and H3K18ac are enriched in promoter regions. The promoter was defined as regions ± 2 kb around known transcription start sites. **m, n**, H3K18la and H3K18ac correlate with steady-state mRNA levels. The average ChIP signal intensity (read count per million mapped reads) for indicated antibodies is shown for genes with different expression levels (the top 25%, the second 25%, the third 25%, and the bottom 25% of RNA-seq counts). **o, p**, IGV tracks for *Arg1* and *Crem* from ChIP–seq analysis, representing data from single experiment. Data in **a**–**f** represent two independent experiments. Data in **g**, **i**–**k** represent three independent experiments.

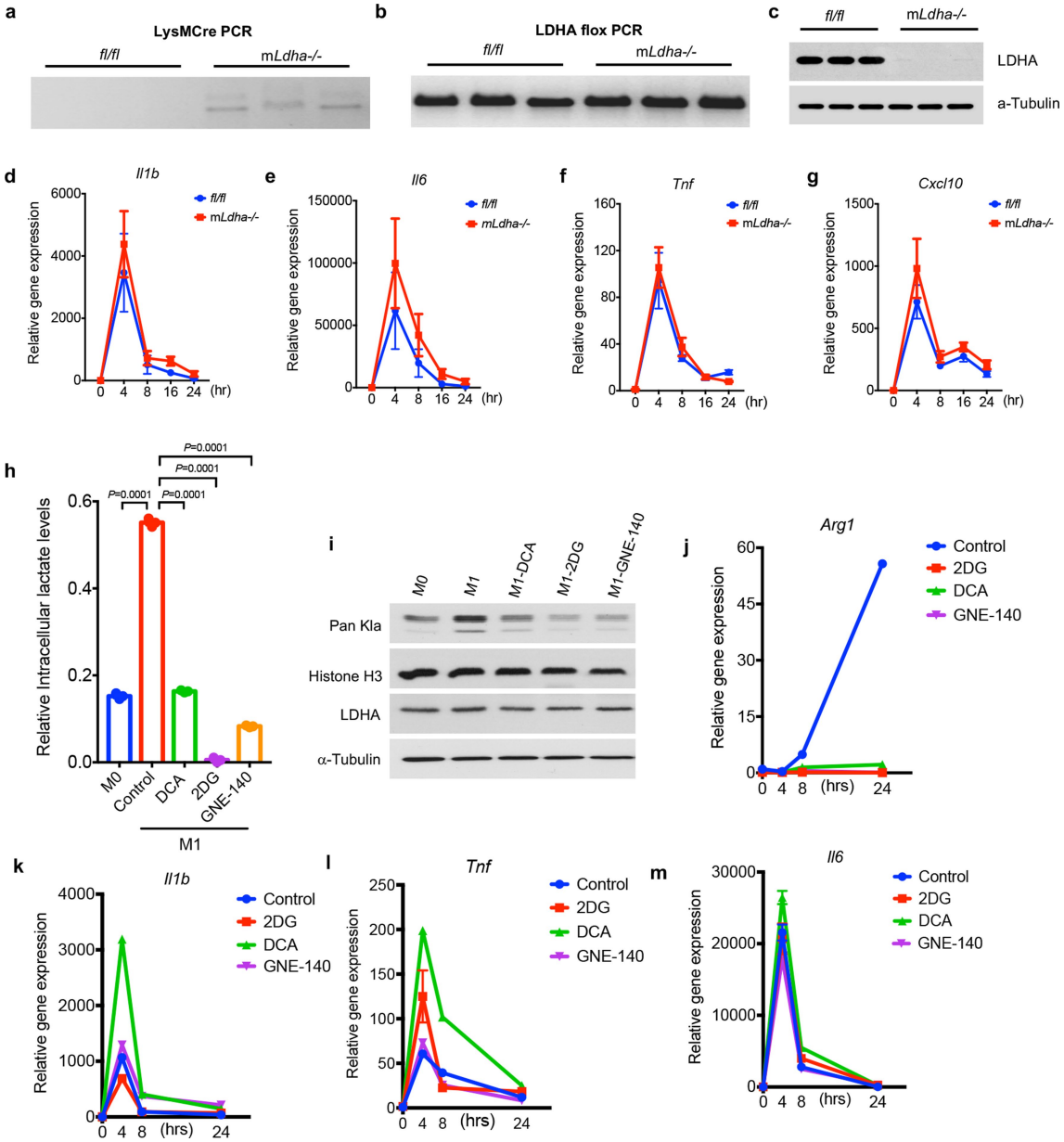



**Extended Data Fig. 5** | See next page for caption.

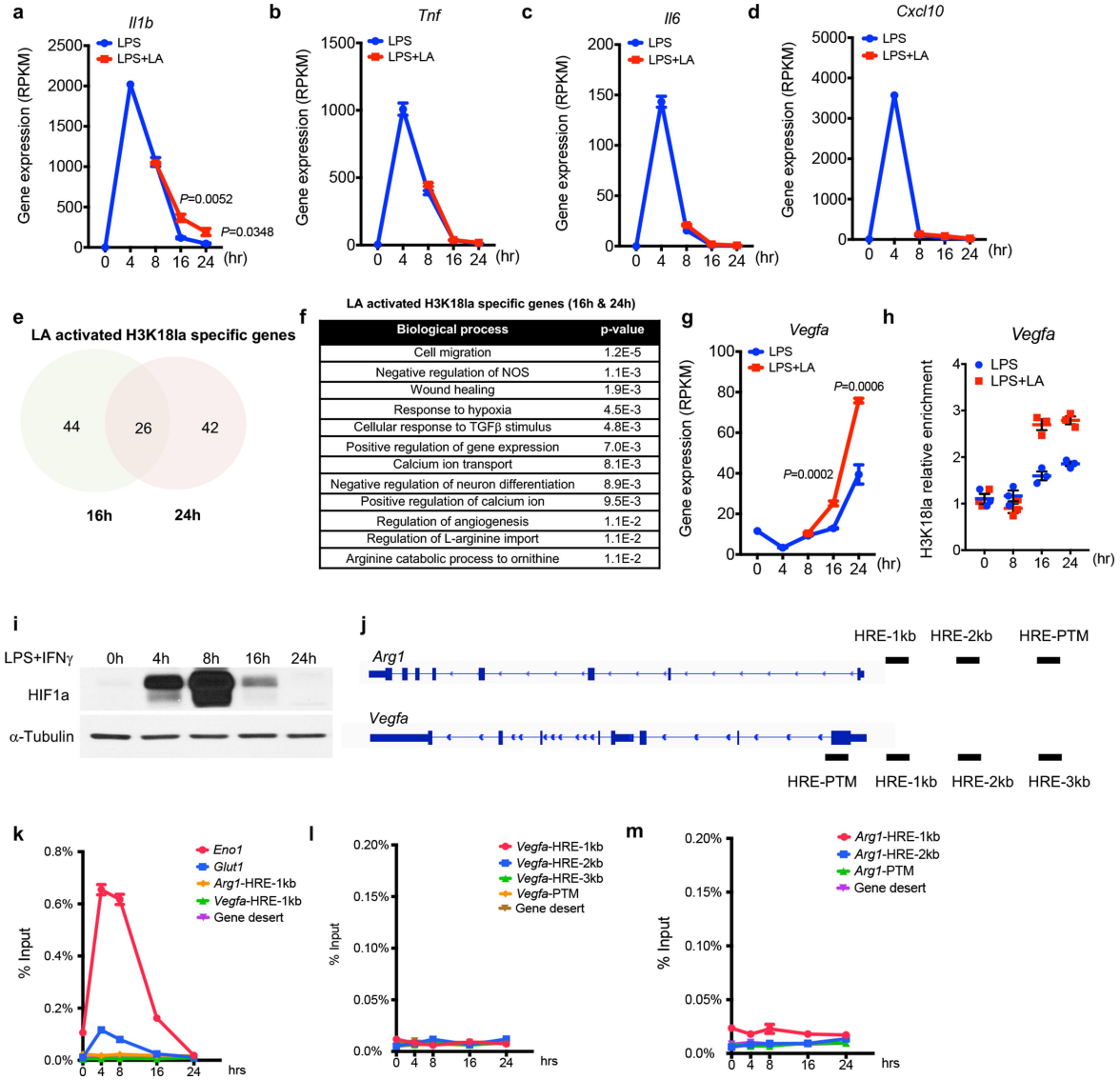
**Extended Data Fig. 5 | Histone Kla-specific genes are associated with late activated M2-like gene expression. a, b**, Heat maps showing expression kinetics of total genes (**a**) and H3K18la-specific genes (**b**) during M1 macrophage polarization. $n = 4$ biological replicates. The colour key represents $\log_2$-transformed fold change relative to the mean of each row. Arrows next to the heatmaps refer to late activated genes (16–24 h) from H3K18la-specific or total genes used for contingency test. **c**, Contingency table analysis (Fisher's exact tests) shows the relation between specific H3K18la enrichment (H3K18la $\log_2$-transformed fold change ≥ 1 and H3K18ac $\log_2$-transformed fold change ≤ 0.5) and late gene activation. **d**, Gene Ontology analysis (biological processes) of H3K18la-specific genes. Statistical significance was determined by modified Fisher's exact test (EASE score) using DAVID bioinformatics resources 6.8; $n = 1{,}223$ genes. **e**–**j**, BMDM cells were infected with indicated Gram-negative bacteria for 24 h. Intracellular lactate (**e**) and histone Kla levels (**f**) were measured 24 h after bacterial challenge. **e**, $n = 3$ biological replicates; statistical significance was determined using one-way ANOVA followed by Dunnett's multiple comparisons test. **g**–**j**, Gene expression was analysed by RT–qPCR at indicated time points after bacterial challenge. **k**, Activities of iNOS and ARG1 were analysed by and commercialized kits from BMDMs activated by the indicated stimuli. Data are mean and s.e.m. from three biological replicates. Data in **f** and **k** represent three independent experiments.



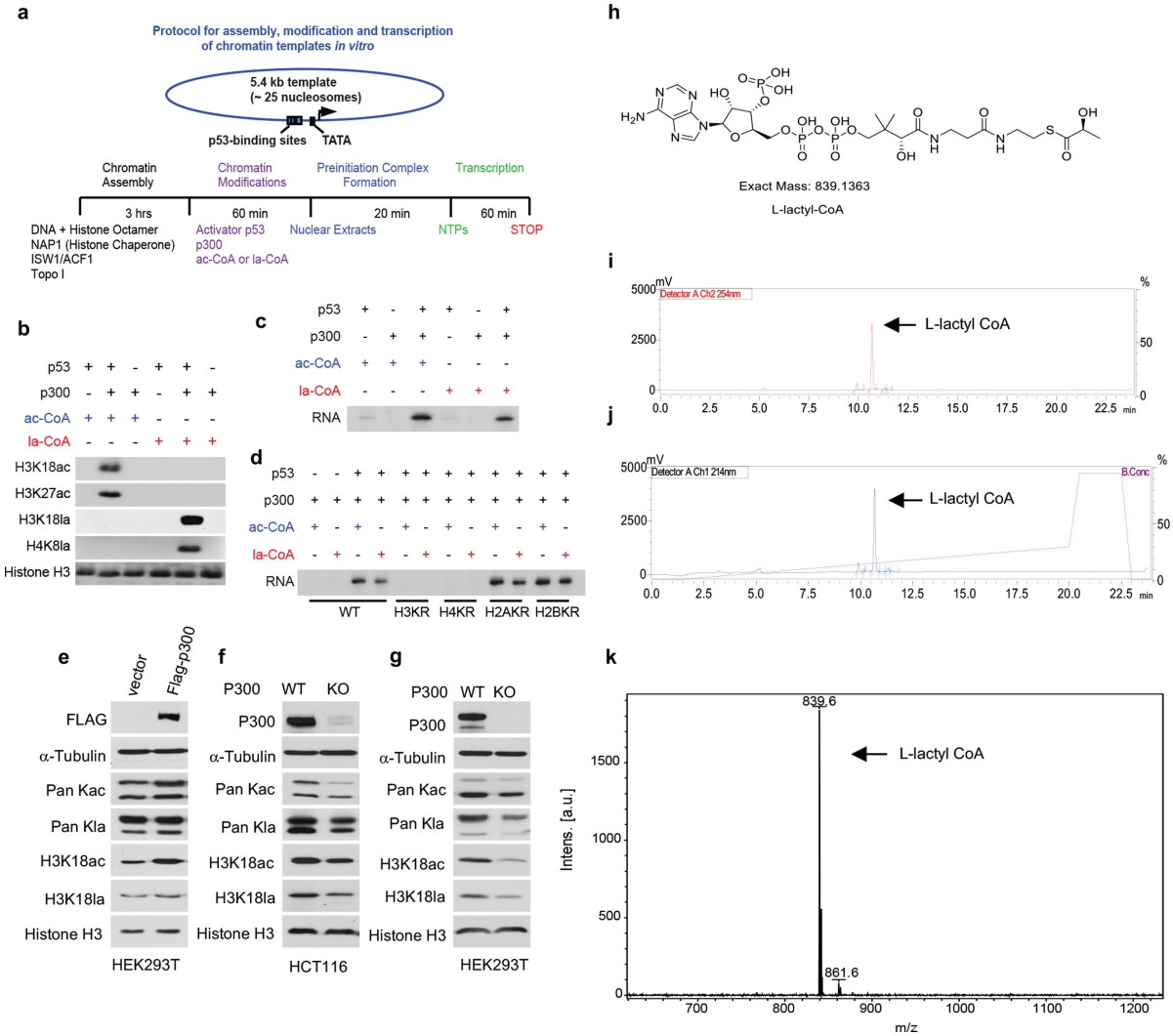

**Extended Data Fig. 6 | Histone Kla levels are positively correlated with *Arg1* expression in TAMs. a**, The purity of TAMs and peritoneal macrophages (PMs) was confirmed by flow cytometry using CD11b and F4/80 antibodies. **b**–**e**, Data were quantified by FlowJo v.10.4.1. Histone Kla and Kac levels were analysed by immunoblots (**b**), intracellular lactate was measured using a lactate colorimetric assay kit (**c**), and gene expression of *Arg1* and *Vegfa* were analysed by RT–qPCR (**d**, **e**) from FACS-sorted peritoneal macrophages and TAMs within the tumour from LLC and B16 tumours. Data in **c**–**e** are mean and s.e.m. $n = 5$ biological independent samples; statistical significance was determined using one-way ANOVA followed by Dunnett's multiple comparisons test. Data in **a** and **b** represent five independent mice.




**Extended Data Fig. 7 | Decreased lactate production lowered histone Kla levels and *Arg1* expression during M1 polarization. a, b**, Genotyping of *Ldha*$^{fl/fl}$ × LysM-Cre$^{+/-}$ mice. **c**, Genotype validation by LDHA immunoblot analysis. **d**–**g**, Gene expression analysis of cytokines by RT–qPCR at indicated time points after M1 polarization. **h**–**m**, Intracellular lactate levels (**h**) were analysed using a lactate colorimetric assay kit and global histone Kla levels (**i**) were measured by immunoblots 24 h after M1 polarization. Inhibitors were treated 30 min after M1 polarization. Gene expression was analysed by RT–qPCR at indicated time points after M1 polarization (**j**–**m**). Data are mean and s.e.m. from three biological replicates. Statistical significance was determined using one-way ANOVA followed by Dunnett's multiple comparisons test. Data in **a**–**c** and **i** represent three independent experiments.




**Extended Data Fig. 8 | Exogenous lactate activates M2-like gene expression through histone Kla. a**–**d**, Exogenous lactate (LA) does not interfere with gene expression of inflammatory cytokines. Data are mean ± s.e.m. from four biological replicates. **e**, Number of lactate-activated H3K18la-specific genes at indicated times are shown in a Venn diagram. **f**, Gene Ontology analysis (biological processes) of lactate-induced H3K18la-specific genes at 16 and 24 h after M1 polarization. Statistical significance was determined by modified Fisher's exact test (EASE score) using DAVID bioinformatics resources 6.8; $n = 112$ genes. **g**, *Vegfa* was induced by exogenous lactate during M1 macrophage polarization; $n = 4$ biological replicates; statistical significance was determined using multiple *t*-tests corrected using the Holm–Sidak method. **h**, H3K18la occupancy at the *Vegfa* promoter was analysed by ChIP–qPCR at indicated time and treatment; data represent three technical replicates from pooled samples. **i**–**m**, HIF1a is not required for histone Kla-mediated *Arg1* induction during M1 polarization. **i**, Immunoblot of HIF1a at indicated time points after M1 polarization. **j**, Illustration of genomic loci targeted by *Arg1* and *Vegfa* ChIP–qPCR primers. HRE indicates regions containing the putative HIF1a binding motif 'ACGTG'. **k**–**m**, ChIP–qPCR analysis of HIF1a binding to indicated genomic locations; data represent three technical replicates from pooled samples. Data are mean and s.e.m. Data in **i** represent three independent experiments.




**Extended Data Fig. 9 | Histone Kla directly stimulates gene transcription from recombinant chromatin in vitro. a**, Protocol for assembly, modification and transcription of chromatin templates. **b**, P300 catalyses histone lactylation in a p53-dependent manner. **c**, Histone lactylation directly stimulates p53-dependent transcription from recombinant chromatin. **d**, H3 and H4 lysine-to-arginine (KR) mutations eliminate p300-dependent transcriptional activation by p53. Recombinant chromatin was assembled with wild-type or H3KR, H4KR, H2AKR or H2BKR mutant histones as indicated. **e**, HEK293T cells were transfected with vector or Flag-tagged p300 plasmid. At 48 h after transfection, whole-cell lysates were prepared and immunoblotted with indicated antibodies. **f**, **g**, Immunoblots of histone Kla and Kac levels in HCT116 (**f**) and HEK293T cells (**g**) in which p300 was **genetically del**eted. **h**–**k**, Quality control of synthesized L-lactyl-CoA. **h**, Illustration of L-lactyl-CoA structure. **i**, **j**, HPLC analysis of the synthesized L-lactyl-CoA. The UV detection wavelength was fixed at 214 and 254 nm. **k**, MALDI-mass spectrometry analysis of L-lactyl-CoA. Data in **b**–**g** and **i**–**k** represent three independent experiments.

nature research

Corresponding author(s): Yingming Zhao; Lev Becker

Last updated by author(s): Aug 15, 2019

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist .

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☒ | ☐ | A description of all covariates tested |
| ☒ | ☐ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ | For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☒ | ☐ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒ | ☐ | Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |

*Our web collection on statistics for biologists contains articles on many of the points above.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | Mass spectrometry: Thermo Xcalibur 3.0.63<br>Flow cytometry: FACSDiva 10.5.3 |
| Data analysis | GraphPad 7.0 was used to perform general statistical analyses.<br>ChIP-seq and RNAseq:<br>FastQC version 0.11.4, Bowtie version 2.2.6, SAMtools version 0.1.19, MACS version 2.2.1, featureCounts version 1.5.0-p1, DAVID Bioinformatics Resources 6.8, HISAT2 version 2.1.0, edgeR version 3.16.5, and Perseus version 1.6.1.1.<br>Mass spectrometry: MaxQuant 1.3.0.5.<br>Flow Cytometry: FlowJo v.10.4.1. |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The ChIP-seq and RNA-seq data have been made available at the Gene Expression Omnibus (GEO) repository under the accession number GSE115354. All other data are available from the authors upon reasonable request.

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Sample size | Specific sample sizes are described in figures or figure legends for all experiments. The sample size used for animal experiment is based on previous experience from the Zhao and Becker labs. No statistical test was used to pre-determine sample size. |
| Data exclusions | No samples or animals were excluded from the analyses. |
| Replication | The number of repeats for each experiments are described in corresponding figure legends. All repeats support the same conclusion. |
| Randomization | Cells or mice tissue were randomly assigned to groups (chemical compound/hypoxia/other treatments). |
| Blinding | For animal related experiments, the investigators were divided into two groups: one group is responsible for collecting samples and the other group is responsible for experiment and outcome assessment. |

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials & experimental systems

| n/a | Involved in the study |
|---|---|
| ☐ | ☒ Antibodies |
| ☐ | ☒ Eukaryotic cell lines |
| ☒ | ☐ Palaeontology |
| ☐ | ☒ Animals and other organisms |
| ☒ | ☐ Human research participants |
| ☒ | ☐ Clinical data |

## Methods

| n/a | Involved in the study |
|---|---|
| ☐ | ☒ ChIP-seq |
| ☐ | ☒ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |

# Antibodies

Antibodies used

The following antibodies were generated by PTM Bio Inc (Chicago, IL):
pan anti-Kac (PTM-101), 1:2000 (WB)
pan anti-Kla (PTM-1401), 1:2000 (WB)
anti-H3K18la (PTM-1406), 1:5000 (WB), 4ug per per ChIP
anti-H4K8la (PTM-1405), 1:5000 (WB)
anti-H4K5la (PTM-1407), 1:5000 (WB)

The following antibodies were generated by Abcam (Cambridge, MA):
anti-histone H3 (ab12079), 1:10000(WB)
anti-H3K18ac (ab1191), 1:10000 (WB), 4ug per ChIP
anti-H3K27ac (ab4729), 1:5000 (WB)

The following antibodies were generated by Active Motif (Carlsbad, CA):
anti-drosophila spike-in antibody (61686), 2ug per ChIP

The following antibodies were generated by Cell Signaling Technology (Danvers, MA):
anti-LDHA (2012S), 1:2000 (WB)

The following antibodies were generated by Millipore Sigma (Burlington, MA):
anti-a-Tubulin (05-829), 1:5000 (WB)
anti-LDHB (ABC927), 1:2000 (WB)

The following antibodies were generated by Novus Biologicals (Littleton, CO):
anti-HIF-1a (NB100-105), 1:2000 (WB)

The following antibodies were generated by GeneTex (Irvine, CA):
anti-iNOS (GTX130246), 1:2000 (WB)
anti-Arg1 (GTX109242), 1:2000 (WB)

The following antibodies were generated by Santa Cruz Biotechnology, Inc (Dallas, TX):
anti-p300 (sc-584), 1:2000 (WB)

The following antibodies were generated by ThermoFisher Scientific (Waltham, MA):
anti-CD11b Monoclonal Antibody (M1/70), PE-Cyanine7, eBioscience (25-0112-82), 0.125 µg/test (Flow)
anti-F4/80 Monoclonal Antibody (BM8), APC, eBioscience (17-4801-82), 2 µg/test (Flow)

The following antibodies were generated by Jackson ImmunoResearch Laboratories (West Grove, PA):
Peroxidase AffiniPure Goat Anti-Mouse IgG (H+L) (115-035-003), 1:10000 (WB)
Peroxidase AffiniPure Goat Anti-Rabbit IgG (H+L) (111-035-003), 1:10000 (WB)

Validation

Pan anti-Kac (PTM-101):
Species: human, mouse; Application: Western Blot, Immunoprecipitation; Manufacturer's web site: https://www.ptmbiolabs.com/product/ptm-101/
Pan anti-Kla (PTM-1401):
Species: human, mouse; Application: Dot Blot, Western Blot, Immunoprecipitation; Validated in this paper.
Anti-H4K8la (PTM-1405):
Species: human, mouse; Application: Western Blot; Validated in this paper.
Anti-H3K18la (PTM-1406):
Species: human, mouse; Application: Dot Blot, Western Blot, ChIP; Validated in this paper.
Anti-H4K5la (PTM-1407):
Species: human, mouse; Application: Dot Blot, Western Blot; Validated in this paper.

Anti-histone H3 (ab12079):
Species: human, mouse; Application: Western Blot; Manufacturer's web site: https://www.abcam.com/histone-h3-antibody-chip-grade-ab12079.html
Anti-H3K18ac (ab1191):
Species: human, mouse; Application: Western Blot, ChIP; Manufacturer's web site: https://www.abcam.com/histone-h3-acetyl-k18-antibody-chip-grade-ab1191.html
Anti-H3K27ac (ab4729):
Species: human, mouse; Application: Western Blot; Manufacturer's web site: https://www.abcam.com/histone-h3-acetyl-k27-antibody-chip-grade-ab4729.html

Anti-LDHA (2012S):
Species: human, mouse; Application: Western Blot; Manufacturer's web site: https://www.cellsignal.com/products/primary-antibodies/ldha-antibody/2012

Anti-a-Tubulin (05-829):
Species: human, mouse; Application: Western Blot; Manufacturer's web site: https://www.emdmillipore.com/US/en/product/Anti-Tubulin-Antibody-clone-DM1A,MM_NF-05-829
Anti-LDHB (ABC927):
Species: human, mouse; Application: Western Blot; Manufacturer's web site: https://www.emdmillipore.com/US/en/product/Anti-LDHB-Antibody,MM_NF-ABC927

Anti-drosophila spike-in antibody (61686):
Species: drosophila; Application: ChIP; Manufacturer's web site: https://www.activemotif.com/catalog/1091/chip-normalization

Anti-iNOS (GTX130246):
Species: mouse; Application: Western Blot; Manufacturer's web site: https://www.genetex.com/Product/Detail/iNOS-antibody/GTX130246
Anti-Arg1(GTX109242):
Species: mouse; Application: Western Blot; Manufacturer's web site: https://www.genetex.com/Product/Detail/Arginase-1-antibody/GTX109242

anti-p300 (N15) (sc-584):
Species: human; Application: Western Blot; Manufacturer's web site: https://www.scbt.com/scbt/product/p300-antibody-n-15
p300 (N-15) has been discontinued and replaced by p300 (F-4): sc-48343.

Anti-CD11b Monoclonal Antibody (M1/70), PE-Cyanine7, eBioscience (25-0112-82):
Species: mouse; Application: Flow cytometry; Manufacturer's web site: https://www.thermofisher.com/antibody/product/CD11b-Antibody-clone-M1-70-Monoclonal/25-0112-82
Anti-F4/80 Monoclonal Antibody (BM8), APC, eBioscience (17-4801-82),
Species: mouse; Application: Flow cytometry; Manufacturer's web site: https://www.thermofisher.com/antibody/product/F4-80-Antibody-clone-BM8-Monoclonal/17-4801-82

## Eukaryotic cell lines

Policy information about cell lines

| | |
|---|---|
| Cell line source(s) | HeLa, MCF-7, A549, MDA-MB-231, HepG2, MEF, and RAW264.7 cells were obtained from the American Type Culture Collection. |
| Authentication | HeLa, MCF-7, A549, MDA-MB-231, HepG2, MEF, and RAW264.7 cells were authenticated based on our vast experience working with these cells lines (such as cell morphology, culture conditions, etc.). Furthermore, we believe that the modification we described in the paper is widely existed in various cell lines, not specific to certain cell types. |
| Mycoplasma contamination | Cells were routinely tested for mycoplasma contamination, and only negative cells were used in experiments |
| Commonly misidentified lines (See ICLAC register) | None of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC. |

## Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

| | |
|---|---|
| Laboratory animals | Adult male mice (Mus musculus, C57BL/6, 7-10 weeks old) were purchased from The Jackson Laboratory, and were used to generate bone marrow derived macrophages (BMDMs). Ldhafl/fl mice (Jackson laboratory, 030112) and LysMcre mice (Jackson laboratory, 004781) were used to generate LysMcre+/- Ldhafl/fl and littermate control LysM-cre-/- Ldhafl/fl mice. |
| Wild animals | The study did not involve samples collected from wild animals. |
| Field-collected samples | The study did not involve samples collected from the field. |
| Ethics oversight | All animal protocols were approved by Institutional Animal Care and Use Committee (ACUP) at the University of Chicago. |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## ChIP-seq

### Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

| | |
|---|---|
| Data access links<br>*May remain private before publication.* | https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE115354 |
| Files in database submission | GSM3176446_RH_148_peaks.broadPeak.gz<br>GSM3176447_RH_229_peaks.broadPeak.gz<br>GSM3176449_RH_151_peaks.broadPeak.gz<br>GSM3176450_RH_231_peaks.broadPeak.gz<br>GSE115354_ChIP-seq_normalization.txt.gz |
| Genome browser session (e.g. UCSC) | N/A |

### Methodology

| | |
|---|---|
| Replicates | ChIP-seq samples are prepared as one replicate, pooled from four mice |
| Sequencing depth | RH_148 (unique reads: 21765076; spike-in reads: 310066)<br>RH_151 (unique reads: 21006731; spike-in reads: 141688)<br>RH_229 (unique reads: 36576826; spike-in reads: 52482)<br>RH_231 (unique reads: 32948615; spike-in reads: 50470) |
| Antibodies | The anti-H3K18la antibody was generated by PTM biolabs. The process for generating antibodies were described similarly in Cell, 2011. 146: p. 1016-1028. Mol Cell, 2015. 58(2): p. 203-15. Nat Chem Biol, 2014. 10(5): p. 365-70. except for using different immunogens.<br>The anti-H3K18ac antibody was purchased from Abcam(ab1191, lot GR 300534-1) |

| | |
|---|---|
| | The evaluation of the antibody for specificity and ChIP grade is provided in the manuscript.<br>Spike-in information: spike-in chromatin (Active motif, Catalog No. 53083), spike-in antibody (Active motif, Catalog No. 61686) |
| Peak calling parameters | Peaks were called using MACS version 2.2.1 under q value = 0.01. |
| Data quality | Sequencing quality was evaluated by FastQC version 0.11.4. All reads were mapped to reference genome of illumina iGenomes UCSC mm10 using Bowtie version 2.2.6, and only uniquely mapped reads were retained. SAMtools version 0.1.1926 was used to convert files to bam format, sort, and remove PCR duplicates. Peaks were called using MACS version 2.2.1 under q value = 0.01.<br>The number of peaks at the cutoff threshold in each sample:<br>21885 peaks in GSM3176446_RH_148_peaks.broadPeak.gz<br>41493 peaks in GSM3176447_RH_229_peaks.broadPeak.gz<br>16139 peaks in GSM3176449_RH_151_peaks.broadPeak.gz<br>42237 peaks in GSM3176450_RH_231_peaks.broadPeak.gz |
| Software | Base called by Real-Time Analysis (RTA)<br>Reads were mapped to reference genome of illumina iGenomes UCSC mm10 using Bowtie version 2.2.6, and only uniquely mapped reads were retained.<br>SAMtools version 0.1.19 was used to convert files to bam format, sort, and remove PCR duplicates.<br>Peaks were called using MACS version 2.2.1 under q value = 0.01.<br>Uniquely mapped reads of each gene were counted by featureCounts version 1.5.0-p1, and normalized by corresponding uniquely mapped spiked-in ChIP read counts. |

# Flow Cytometry

## Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

## Methodology

| | |
|---|---|
| Sample preparation | 0.2 million cells were labeling with different fluorophore conjugated antibody at room temperature for 15mins, followed by two washes. |
| Instrument | Samples were analyzed using a FACSCanto™ II flow cytometer. |
| Software | Data were quantified by FlowJo v.10.4.1. |
| Cell population abundance | Purity for both TAM and Pmac are above 95% based on F4/80 and CD11b cell surface marker. positive gating were determined by negative control. |
| Gating strategy | Cells were gated by FSC/SSC for total population --> SSC-A/SSC-H for single cells --> cblue labeling for live cells population --> F4/80 and CD11b double positive (compared to negative population) for purity check. |

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Article

# Structural basis of species-selective antagonist binding to the succinate receptor

https://doi.org/10.1038/s41586-019-1663-8

Received: 11 December 2018

Accepted: 17 September 2019

Published online: 23 October 2019

Matthias Haffke[1,2*], Dominique Fehlmann[3], Gabriele Rummel[1], Jacques Boivineau[1], Myriam Duckely[1], Nina Gommermann[2], Simona Cotesta[2], Finton Sirockin[2], Felix Freuler[1], Amanda Littlewood-Evans[3], Klemens Kaupmann[3*] & Veli-Pekka Jaakola[1,4*]

The tricarboxylic acid cycle intermediate succinate is involved in metabolic processes and plays a crucial role in the homeostasis of mitochondrial reactive oxygen species[1]. The receptor responsible for succinate signalling, SUCNR1 (also known as GPR91), is a member of the G-protein-coupled-receptor family[2] and links succinate signalling to renin-induced hypertension, retinal angiogenesis and inflammation[3–5]. Because SUCNR1 senses succinate as an immunological danger signal[6]—which has relevance for diseases including ulcerative colitis, liver fibrosis[7], diabetes and rheumatoid arthritis[3,8]—it is of interest as a therapeutic target. Here we report the high-resolution crystal structure of rat SUCNR1 in complex with an intracellular binding nanobody in the inactive conformation. Structure-based mutagenesis and radioligand-binding studies, in conjunction with molecular modelling, identified key residues for species-selective antagonist binding and enabled the determination of the high-resolution crystal structure of a humanized rat SUCNR1 in complex with a high-affinity, human-selective antagonist denoted NF-56-EJ40. We anticipate that these structural insights into the architecture of the succinate receptor and its antagonist selectivity will enable structure-based drug discovery and will further help to elucidate the function of SUCNR1 in vitro and in vivo.

Under certain conditions—including ischaemia reperfusion, hypoxia and in classically activated macrophages—there is an increase in intracellular succinate levels in the mitochondria and cytosol[9]. Ischaemia reperfusion induces succinate dehydrogenase to catalyse its reverse reaction; the subsequent increase in mitochondrial succinate then drives—through reversal of the electron-transport chain—the production of reactive oxygen species and, ultimately, tissue damage[10]. In classically activated glycolytic macrophages, increased levels of cytoplasmic succinate lead to stabilization of HIF1α and enhanced production of pro-inflammatory factors through glutamine-dependent anaplerosis and the 'GABA (γ-aminobutyric acid) shunt' pathway[11,12]. The role of SUCNR1 in these intracellular processes is unknown; however, hypoxia, necrosis and inflammation also result in the extracellular accumulation of succinate, where it triggers SUCNR1 signalling[13]. The succinate–SUCNR1 axis drives a range of events in many tissues—including eye, kidney, liver and gut[13] (Extended Data Fig. 1)—and has been associated with metabolic indications[3,13]. Conversely, however, SUCNR1 signalling in myeloid cells is reported to resolve acute inflammation in obesity[14].

Small non-metabolite agonists of SUCNR1 have been identified using molecular modelling and docking based on structures of the P2Y1 receptor[15]. High-affinity and selective antagonists have also been reported[16], but structural details of the interaction between antagonists and SUCNR1 remain unknown. Despite extensive mutagenesis studies investigating the species-selectivity of SUCNR1 agonists[15], it is not clear how the differences in structure–activity relationships arise among the SUCNR1 orthologues.

To enable structure determination, we screened human and rat SUCNR1 orthologues for optimal expression, purification and stabilization, as well as screening various fusion proteins and detergents typically used to stabilize G-protein-coupled receptors (GPCRs)[17]. We selected the rat orthologue on the basis of its lack of any glycosylation site and its increased biochemical stability in comparison with human SUCNR1.

We generated nanobodies to stabilize and trap the receptor in a single conformation, according to established protocols[18]. Nanobody6 increased the thermal stability of wild-type rat SUCNR1, formed a stable complex with the receptor and acted as a negative allosteric modulator in a [$^{35}$S]GTPγS assay, and could thus potentially stabilize an inactive conformation of the receptor (Extended Data Fig. 2a–d). Nanobody6 also formed a stable complex with human SUCNR1, and showed a similar—albeit much weaker—modulation of receptor activity (Extended Data Fig. 2e, f). The unusually high thermal stability of rat SUCNR1, in conjunction with the conformational stability provided by Nanobody6, enabled us to crystallize the full-length wild-type rat SUCNR1–Nanobody6 complex. We obtained crystals in lipidic cubic phase. The crystals diffracted anisotropically to a resolution of 2.1 Å (Extended Data Fig. 3a, b, Extended Data Table 1), and the resulting electron density maps were of excellent quality (Extended Data Fig. 3c, d).

[1]Chemical Biology & Therapeutics, Novartis Institutes for BioMedical Research, Novartis Pharma AG, Basel, Switzerland. [2]Global Discovery Chemistry, Novartis Institutes for BioMedical Research, Novartis Pharma AG, Basel, Switzerland. [3]Autoimmunity, Transplantation and Inflammation, Novartis Institutes for BioMedical Research, Novartis Pharma AG, Basel, Switzerland. [4]Present address: Confo Therapeutics, Zwijnaarde, Belgium. *e-mail: matthias.haffke@novartis.com; klemens.kaupmann@novartis.com; veli-pekka.jaakola@confotherapeutics.com

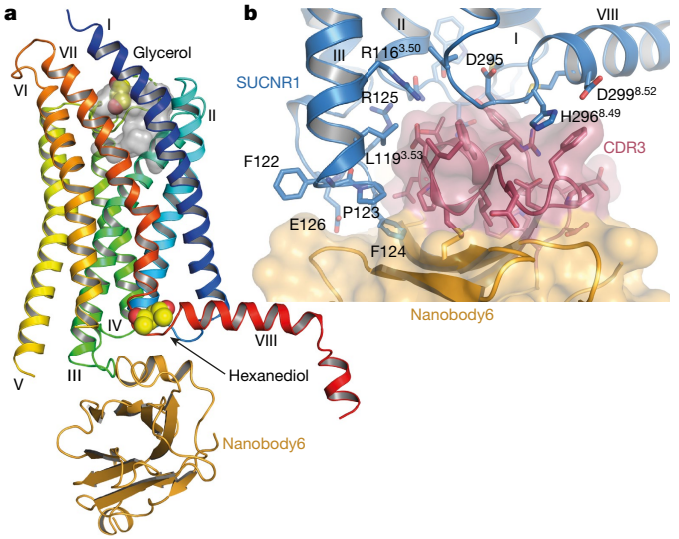



**Fig. 1 | Structure of the apo rat SUCNR1–Nanobody6 complex. a**, Side view of the rat SUCNR1–Nanobody6 complex. SUCNR1 is coloured in blue to red from N- to C terminus. Glycerol and 2,5-hexanediol, identified in electron-density maps, are shown as sphere models. The potential orthosteric ligand-binding site is shown as a grey surface. **b**, Detailed view of the SUCNR1–Nanobody6 interface. The CDR3 of Nanobody6 is shown in red and side chains of important residues are shown as sticks.

Our structure of the rat SUCNR1–Nanobody6 complex involves the receptor in an inactive conformation, based on comparisons with structures of the β2-adrenergic receptor (β2-AR)[19,20] and the P2Y1 receptor[21] (Extended Data Fig. 4a, b). Nanobody6 binds to the intracellular side of the receptor (Fig. 1a), at an interface that is similar to but distinct from that of the Gα subunit of the G-protein trimer (Extended Data Fig. 4c, d), and the interaction surface area is around 750 Å$^2$. In comparison with other GPCR–nanobody complexes[22–26], Nanobody6 has an unusually long complementarity-determining region 3 (CDR3) (Extended Data Fig. 4e, f). The interface between SUCNR1 and Nanobody6 is stabilized by several hydrogen bonds and hydrophobic interactions, involving CDR3 of Nanobody6 and intracellular loop 1 (ICL1), ICL2, ICL4 and helix VIII in SUCNR1 (Fig. 1b). The binding site for Nanobody6 on the receptor is highly conserved between rat, human and mouse SUCNR1 (Extended Data Fig. 5a, b). In view of its overlapping binding site with Gα, Nanobody6 is expected to affect both $G_{i/o}$- and $G_q$-mediated SUCNR1 signalling.

We purified and crystallized the receptor–nanobody complex in the presence of a previously reported antagonist (compound 5g)[16]; however, we did not observe any electron density in our structure for this compound. Possible explanations for the missing ligand density could be the low solubility of compound 5g, potential competition for binding with Nanobody6, or absorption into the lipidic cubic phase. We modelled two positions with unexplained electron density as glycerol and 2,5-hexanediol, respectively (Extended Data Fig. 3e–j). Although both influence the stability of the receptor, we could not attribute any functional effects to these molecules, and so categorize them as crystallization artefacts (Extended Data Fig. 2g–k). The glycerol molecule is located at a positively charged entry site to a hydrophobic pocket, which is surrounded by helix I, helix II, extracellular loop 1 (ECL1), helix III, helix VI and helix VII, and is partially occluded from solvent by ECL2. The lower part of this pocket is highly conserved, and is faced by side chains that were previously reported to be important for receptor activation by succinate[2] (R95$^{3.29}$, H99$^{3.33}$, R248$^{6.55}$ and R276$^{7.39}$; superscripts denote Ballesteros–Weinstein numbering for GPCRs[27]; Extended Data Fig. 6a, b). H99$^{3.33}$ and R248$^{6.55}$ constitute an intricate interhelical hydrogen-bond network with Y103$^{3.37}$ and Y244$^{6.51}$; this network stabilizes the overall receptor structure and the lower region of the hydrophobic pocket. R95$^{3.29}$ and S180$^{5.35}$ are involved in the positioning of ECL2 by binding to the backbone carbonyl of V169 and the hydroxyl group of Y171, respectively. The highly conserved residue D170 further stabilizes ECL2 by forming a hydrogen bond with Y272$^{7.35}$ (Extended Data Fig. 6c). These residues could be important for stabilizing the interhelical receptor architecture and for retaining ECL2 close to the orthosteric ligand-binding site. The precise positioning of ECL2 might therefore be a requirement for the formation of a functional succinate-binding site. It has previously been suggested that ECL2 in SUCNR1 might function as a tethered inverse agonist[15] held in place by R95$^{3.29}$; however, this seems unlikely when considering the molecular architecture around ECL2. Our structure of the receptor in the inactive conformation does not enable us to draw conclusions as to how SUCNR1 achieves selectivity for succinate over other very similar dicarboxylates.

We performed a high-throughput screen to identify new SUCNR1 antagonists. 3-(4′-chloro-[1,1′-biphenyl]-3-carboxamido)-3-(pyridin-3-yl)propanoic acid (JC-59-GF68) was identified as a hit, with moderate activity in a [$^{35}$S]GTPγS assay with human SUCNR1. It was optimized via the intermediate 2-(2-(4′-chloro-[1,1′-biphenyl]-3-carboxamido)phenyl)acetic acid (PB-20-OV24) to the high potency antagonist 2-(2-(4′-((4-methylpiperazin-1-yl)methyl)-[1,1′-biphenyl]-3-carboxamido)phenyl)acetic acid (NF-56-EJ40) (Fig. 2a–c, Supplementary Methods). NF-56-EJ40 was highly selective for human SUCNR1 and showed almost no activity towards rat SUCNR1. We then established a radioligand-binding assay for human SUCNR1 using [$^3$H]-labelled NF-56-EJ40 (dissociation constant, $K_d$ = 33 nM; Fig. 2d, e), and investigated the potential orthosteric ligand-binding site of human SUCNR1 by mutagenesis. Residues for mutagenesis were selected on the basis of our structure of rat SUCNR1, differences in primary sequence between rat and human SUCNR1 orthologues (Extended Data Fig. 5) and molecular-docking analysis of NF-56-EJ40 binding to a human SUCNR1 model (Extended Data Fig. 7a). Molecular-docking analysis suggested that NF-56-EJ40 is bound deep inside the hydrophobic pocket, with the acid group coordinated by the hydroxyl groups of the conserved residues Y83$^{2.64}$ and Y30$^{1.39}$ on one side, and R281$^{7.39}$ on the other side. The conserved E18$^{1.27}$ was predicted to form an additional hydrogen bond to the piperazine ring of NF-56-EJ40. E22$^{1.31}$ and N274$^{7.32}$ in human SUCNR1 are replaced by K18$^{1.31}$ and K269$^{7.32}$ in rat SUCNR1. These two amino acid exchanges could prevent the binding of NF-56-EJ40 to rat SUCNR1 owing to steric hindrance, providing a possible rationale for the observed species selectivity. Radioligand-binding studies with human SUCNR1 (Extended Data Table 2, Extended Data Fig. 7b) showed partial agreement with our homology model: the Y30$^{1.39}$F mutant of human SUCNR1 (in which the tyrosine at position 30$^{1.39}$ was mutated to phenylalanine), which was expected to disrupt coordination of the acid moiety, showed reduced binding of NF-56-EJ40. Similar effects were observed with the E18$^{1.27}$K and E18$^{1.27}$R mutants, probably owing to steric clashes of the Lys and Arg residues with NF-56-EJ40 and the loss of a hydrogen bond to its piperazine ring. We next introduced the corresponding rat SUCNR1 residues into human SUCNR1 by preparing the E22$^{1.31}$K/N274$^{7.32}$K double mutant. In this case, we did not obtain a reliable inhibitory constant ($K_i$) owing to a low extent of radioligand binding. However, the reverse case, in which human SUCNR1 residues were introduced into rat SUCNR1 to form the double mutant K18$^{1.31}$E/K269$^{7.32}$N (hereafter denoted humanized rat SUCNR1), bound NF-56-EJ40 with an affinity almost equal to that observed for human SUCNR1 ($K_i$ = 17.4 ± 2.5 nM and $K_i$ = 33.5 ± 7.0 nM for human and humanized rat SUCNR1, respectively; Fig. 2f). To further confirm our observations, we purified humanized rat SUCNR1 and found that NF-56-EJ40 increased the thermal stability of both humanized rat SUCNR1 and human SUCNR1, but not that of rat SUCNR1 (Fig. 2g–i). Taken together, our studies identified E22$^{1.31}$ and N274$^{7.32}$ as key determinants for the species selectivity of NF-56-EJ40 binding.

On the basis of these data, we purified, crystallized and determined the structure of the humanized rat SUCNR1–Nanobody6–NF-56-EJ40 complex (Extended Data Fig. 8a, b). The crystals diffracted anisotropically

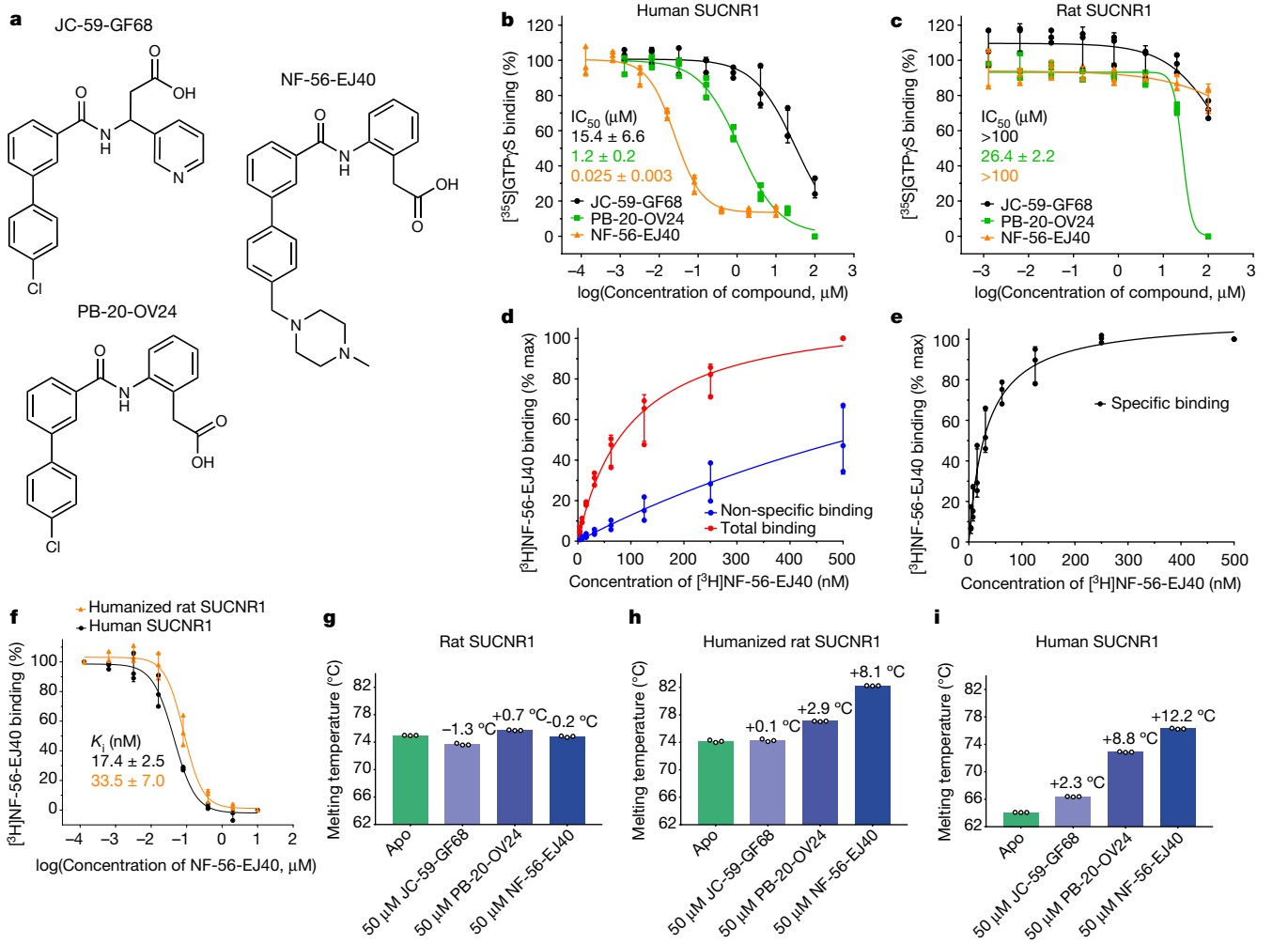



**Fig. 2 | Development of a human SUCNR1 antagonist and humanization of rat SUCNR1. a**, Chemical structures of SUCNR1 antagonists. **b**, **c**, [$^{35}$S]GTPγS assay with human SUCNR1 (**b**) and rat SUCNR1 (**c**) in the presence of 50 μM succinate. **d**, Total and non-specific binding of the radioligand [$^{3}$H]NF-56-EJ40 to human SUCNR1, determined in the presence of 10 mM succinate. **e**, Specific binding of [$^{3}$H]NF-56-EJ40 calculated from **d** (total binding − non-specific binding); $K_d$ = 33 nM. **f**, Radioligand competition binding experiment with wild-type human and humanized rat SUCNR1. For **b**–**f**, curves were calculated from $n$ = 3 independent experiments (data are mean ± s.d.; individual data points are shown). **g**–**i**, Thermal stability assays by nano-differential scanning fluorimetry. Bars represent average melting temperature ($T_m$), and the data points from $n$ = 3 technical replicates are shown as individual circles. $\Delta T_m$ values compared to the control are indicated. The experiment was repeated independently twice with similar results.

to a resolution of less than 2 Å (Extended Data Table 1), and the resulting electron density enabled the unambiguous placement of NF-56-EJ40 in the orthosteric ligand-binding site (Fig. 3a, b, Extended Data Fig. 8c–e). Notably, the binding mode of NF-56-EJ40 was considerably different from that suggested by our molecular-docking studies, and binding was accompanied by large structural rearrangements in helix IV, ECL1 and in particular ECL2 close to the ligand-binding pocket (Extended Data Fig. 8f, g, i–k). Although the acid moiety is in a similar position to that predicted, the overall binding location of NF-56-EJ40 is about 6 Å deeper with respect to the biphenyl and piperazine groups. As suggested by mutagenesis data, the K18$^{1.31}$E mutation is critical for ligand binding because it restores a key receptor–ligand interaction (Extended

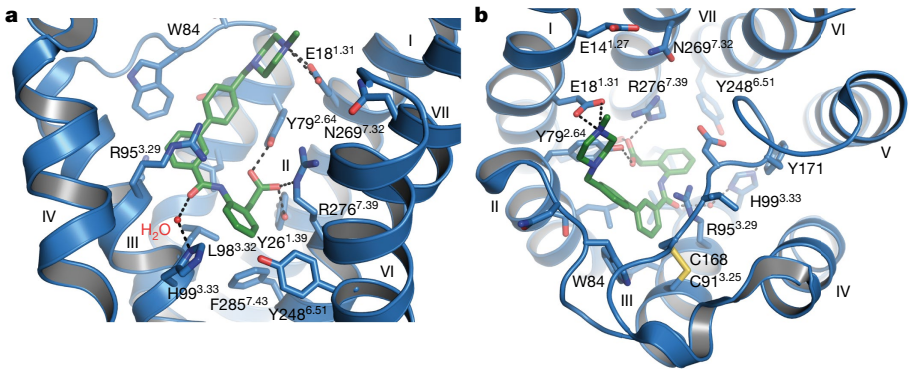



**Fig. 3 | Binding mode of the antagonist NF-56-EJ40 to humanized rat SUCNR1. a**, **b**, Humanized rat SUCNR1 (blue) bound to NF-56-EJ40 (green) in side view (**a**) and top view (**b**). For clarity, ECL2, helix IV and helix V are omitted in the side view. Only key residues within 4 Å of NF-56-EJ40 are shown (as sticks); dashed black lines represent hydrogen bonds. Note the water-mediated hydrogen bond between H99$^{3.33}$ and the amide oxygen in NF-56-EJ40.

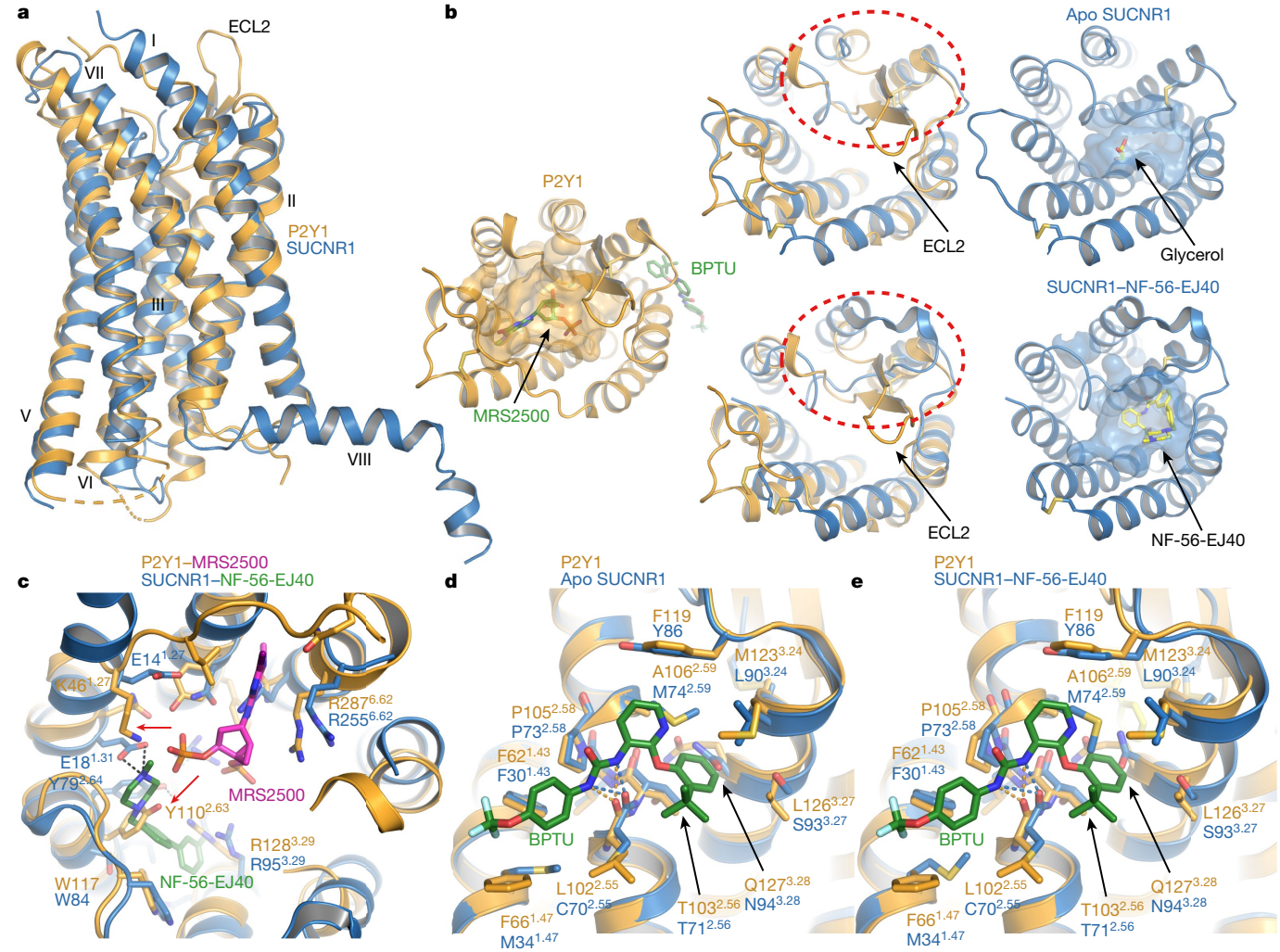



**Fig. 4 | Structural similarities between SUCNR1 and the P2Y1 receptor. a**, Side view of apo rat SUCNR1 and the P2Y1 receptor (PDB ID: 4XNW). P2Y1, orange; SUCNR1, blue. **b**, Top view of aligned receptors. Respective ligands (MRS2500 and BPTU for P2Y1; glycerol or NF-56-EJ40 for SUCNR1) are shown as green or yellow sticks, respectively, with the orthosteric ligand-binding pockets in surface representation. Dashed red circles denote large structural differences between apo rat SUCNR1 and P2Y1 or NF-56-EJ4-bound humanized rat SUCNR1 and P2Y1. **c**, Structural differences in the orthosteric ligand-binding site between NF-56-EJ40-bound humanized rat SUCNR1 and P2Y1 bound to MRS2500 (PDB ID: 4XNW). Red arrows indicate amino acid side chains in P2Y1 that clash with NF-56-EJ40 ($K46^{1.27}$, $Y110^{2.63}$). **d**, **e**, Comparison of the BPTU-binding sites of P2Y1 and apo rat SUCNR1 (**d**) and of P2Y1 and NF-56-EJ40-bound humanized rat SUCNR1 (**e**). BPTU is shown as green sticks. Key residues within 4 Å of BPTU are shown for both receptors. Dashed lines show the conserved binding mode, via hydrogen bonds, to the carbonyl of $L102^{2.55}$ in P2Y1 and $C70^{2.55}$ in rat SUCNR1.

Data Fig. 8g, j). By contrast, the steric hindrance that was predicted for $K269^{7.32}$ was not observed. Binding of the acid moiety of NF-56-EJ40 is mediated via the hydroxyl groups of $Y79^{2.64}$ on one side and by $Y26^{1.39}$ and $R276^{7.39}$ on the other side—a different coordination environment than that predicted. Additional cation–π interactions with $R95^{3.29}$ and van der Waals interactions with $L75^{2.60}$, W84 and $L98^{3.32}$ complete the coordination of NF-56-EJ40, and provide a complex and multivalent ligand-binding mode (Fig. 3a, b, Extended Data Fig. 8h).

SUCNR1 has a high structural similarity to the P2Y1 receptor (root mean square deviation of 1.4 Å to P2Y1; Protein Data Bank (PDB) ID: 4XNW), as has been suggested previously[3] (Fig. 4a). Similar to P2Y1, SUCNR1 contains two disulfide bridges, which position ECL2 close to the orthosteric ligand-binding site ($C91^{3.25}$ to C168) and stabilize the structure between the N terminus and TM7 of the receptor (C7 to $C263^{7.26}$). In P2Y1, binding of the antagonist MRS2500 involves the β-hairpin formed by ECL2, and thus results in a partially closed binding pocket[21]. Our SUCNR1 structures, by contrast, show a rather open extracellular side, in which ECL2 is partially unstructured in the ligand-free form and is noticeably outward-shifted in the antagonist-bound form (Fig. 4a, b). Notably, although the orthosteric ligand-binding sites in both receptors differ substantially (Fig. 4c), we find that some of the key structural features that are important for binding of the allosteric antagonist and antiplatelet agent 1-[2-(2-*tert*-butylphenoxy)pyridin-3-yl]-3-[4-(trifluoromethoxy) phenyl]urea (BPTU) to P2Y1 are also present in SUCNR1. Similar to P2Y1, $P73^{2.58}$ in SUCNR1 precludes intrahelical hydrogen bonding, and the carbonyl of $C70^{2.55}$ is thus available for interaction with the nitrogen atoms of the urea group in BPTU. However, in contrast to the mainly hydrophobic environment in P2Y1 (formed by $T103^{2.56}$, $M123^{3.24}$, $L126^{3.27}$ and $Q127^{3.28}$) that accommodates the aryl group of BPTU, the residues $S93^{3.27}$ and $N94^{3.28}$ in SUCNR1 would be available for hydrogen bonds and thus would favour electrostatic interactions (Fig. 4d, e). These structural similarities between P2Y1 and SUCNR1 therefore highlight the potential to develop allosteric SUCNR1 antagonists based on a BPTU scaffold.

These high-resolution crystal structures of SUCNR1 are, to our knowledge, the first structures of an alicarboxylic acid metabolite receptor. We anticipate that these structures, in conjunction with a newly generated nanobody and a high-affinity antagonist radioligand, will advance the characterization of SUCNR1 in metabolic and immunological disease settings.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1663-8.


1. Tretter, L., Patocs, A. & Chinopoulos, C. Succinate, an intermediate in metabolism, signal transduction, ROS, hypoxia, and tumorigenesis. *Biochim. Biophys. Acta* **1857**, 1086–1101 (2016).
2. He, W. et al. Citric acid cycle intermediates as ligands for orphan G-protein-coupled receptors. *Nature* **429**, 188–193 (2004).
3. Gilissen, J., Jouret, F., Pirotte, B. & Hanson, J. Insight into SUCNR1 (GPR91) structure and function. *Pharmacol. Ther.* **159**, 56–65 (2016).
4. Peruzzotti-Jametti, L. et al. Macrophage-derived extracellular succinate licenses neural stem cells to suppress chronic neuroinflammation. *Cell Stem Cell* **22**, 355–368. e13 (2018).
5. Schneider, C. et al. A metabolite-triggered tuft cell-ILC2 circuit drives small intestinal remodeling. *Cell* **174**, 271–284.e14 (2018).
6. Rubic, T. et al. Triggering the succinate receptor GPR91 on dendritic cells enhances immunity. *Nat. Immunol.* **9**, 1261–1269 (2008).
7. Cho, E. H. Succinate as a regulator of hepatic stellate cells in liver fibrosis. *Front. Endocrinol.* **9**, 455 (2018).
8. Littlewood-Evans, A. et al. GPR91 senses extracellular succinate released from inflammatory macrophages and exacerbates rheumatoid arthritis. *J. Exp. Med.* **213**, 1655–1662 (2016).
9. Jha, A. K. et al. Network integration of parallel metabolic and transcriptional data reveals metabolic modules that regulate macrophage polarization. *Immunity* **42**, 419–430 (2015).
10. Chouchani, E. T. et al. Ischaemic accumulation of succinate controls reperfusion injury through mitochondrial ROS. *Nature* **515**, 431–435 (2014).
11. Tannahill, G. M. et al. Succinate is an inflammatory signal that induces IL-1β through HIF-1α. *Nature* **496**, 238–242 (2013).
12. Mills, E. L. et al. Succinate dehydrogenase supports metabolic repurposing of mitochondria to drive inflammatory macrophages. *Cell* **167**, 457–470 (2016).
13. de Castro Fonseca, M., Aguiar, C. J., da Rocha Franco, J. A., Gingold, R. N. & Leite, M. F. GPR91: expanding the frontiers of Krebs cycle intermediates. *Cell Commun. Signal.* **14**, 3 (2016).
14. Keiran, N. et al. SUCNR1 controls an anti-inflammatory program in macrophages to regulate the metabolic response to obesity. *Nat. Immunol.* **20**, 581–592 (2019).
15. Trauelsen, M. et al. Receptor structure-based discovery of non-metabolite agonists for the succinate receptor GPR91. *Mol. Metab.* **6**, 1585–1596 (2017).
16. Bhuniya, D. et al. Discovery of a potent and selective small molecule hGPR91 antagonist. *Bioorg. Med. Chem. Lett.* **21**, 3596–3602 (2011).
17. Chun, E. et al. Fusion partner toolchest for the stabilization and crystallization of G protein-coupled receptors. *Structure* **20**, 967–976 (2012).
18. Pardon, E. et al. A general protocol for the generation of Nanobodies for structural biology. *Nat. Protoc.* **9**, 674–693 (2014).
19. Cherezov, V. et al. High-resolution crystal structure of an engineered human $\beta_2$-adrenergic G protein-coupled receptor. *Science* **318**, 1258–1265 (2007).
20. Rasmussen, S. G. et al. Crystal structure of the $\beta_2$ adrenergic receptor–Gs protein complex. *Nature* **477**, 549–555 (2011).
21. Zhang, D. et al. Two disparate ligand-binding sites in the human P2Y1 receptor. *Nature* **520**, 317–321 (2015).
22. Che, T. et al. Structure of the nanobody-stabilized active state of the kappa opioid receptor. *Cell* **172**, 55–67 (2018).
23. Staus, D. P. et al. Allosteric nanobodies reveal the dynamic range and diverse mechanisms of G-protein-coupled receptor activation. *Nature* **535**, 448–452 (2016).
24. Burg, J. S. et al. Structural basis for chemokine recognition and activation of a viral G protein-coupled receptor. *Science* **347**, 1113–1117 (2015).
25. Kruse, A. C. et al. Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature* **504**, 101–106 (2013).
26. Rasmussen, S. G. et al. Structure of a nanobody-stabilized active state of the $\beta_2$ adrenoceptor. *Nature* **469**, 175–180 (2011).
27. Ballesteros, J. A. & Weinstein, H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *Methods Neurosci.* **25**, 366–428 (1995).




# Methods

### Data reporting

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

### Nanobody6

The sequence encoding Nanobody6 was cloned into a pIEx/Bac-3-derived plasmid with a haemagglutinin signal sequence followed by a Flag tag at the N terminus of the nanobody and a 10×His-tag preceded by a human rhinovirus (HRV) 3C protease cleavage site at the C terminus.

High-titre recombinant baculovirus was generated with the FlashBac Gold system according to the manufacturer's instructions and used to infect *Spodoptera frugiperda* (Sf9) insect cells at a density of $2.0 \times 10^6$ cells per ml and incubated for 5 days at 27 °C in shaker flasks. The supernatant was collected by centrifugation at 10,000*g* for 20 min at 4 °C, adjusted to 25 mM Tris/HCl pH 8.0, 300 mM NaCl, 5 mM $CaCl_2$ and 1 mM $CoCl_2$ and incubated at 4 °C for 30 min. The precipitant was removed by filtration with glass-fibre pre-filters on top of 0.45-µm Stericup filter devices (Millipore) and the supernatant was subsequently concentrated 8 times with a tangential flow filtration device (Millipore) with a 10 kDa molecular weight cut-off (MWCO) membrane.

Talon resin (Clontech), equilibrated in wash buffer (25 mM HEPES pH 8.0, 300 mM NaCl, 10 mM imidazole) was added to the supernatant and incubated on a shaker at 4 °C for 1 h. The resin was collected by centrifugation at 1,500*g* for 20 min, transferred to a gravity flow column and washed with 30 column volumes of wash buffer before eluting the protein with 5 column volumes of elution buffer (25 mM HEPES pH 8.0, 300 mM NaCl, 250 mM imidazole). His-tagged HRV 3C protease (prepared in-house) was added at a 1:50 (w/w) ratio to cleave the C-terminal 10×His-tag and the sample was dialysed against 20-times excess of dialysis buffer (25 mM HEPES pH 8.0, 300 mM NaCl) at 4 °C overnight. Cleaved Nanobody6 was passed through Ni-NTA resin in a gravity flow column (equilibrated in dialysis buffer) to remove His-tagged HRV 3C protease, uncleaved Nanobody6 and the free 10×His-tag. The flow through was concentrated to a final volume of less than 4 ml and further purified by size-exclusion chromatography on a Superdex S75 16/60 column (GE Healthcare), equilibrated in dialysis buffer. Protein containing fractions were pooled, concentrated to around 7 mg $ml^{-1}$ with a 10 kDa MWCO concentrator (Millipore), flash-frozen in small aliquots in liquid nitrogen and stored at −80 °C until use. Final yields were about 4 mg of protein per litre of expression culture. Mass spectrometry confirmed the presence of two disulfide bridges in Nanobody6.

### Rat SUCNR1

Wild-type rat SUCNR1 (residues 2 to 317) was cloned into a pIEx/Bac-3-derived plasmid with a haemagglutinin signal sequence followed by a Flag tag at the N terminus of the receptor and a 10×His-tag preceded by a HRV 3C protease cleavage site at the C terminus.

High-titre recombinant baculovirus was generated with the FlashBac Gold system according to the manufacturer's instructions and used to infect *Spodoptera frugiperda* (Sf9) insect cells at a density of $2.0 \times 10^6$ cells per ml and incubated for 3 days at 27 °C in shaker flasks. The cells were collected by centrifugation at 10,000*g* for 20 min at 4 °C, frozen and stored at −20 °C until further use. Membranes were prepared by lysing cells in hypotonic buffer (10 mM HEPES, pH 7.5, 10 mM $MgCl_2$, 20 mM KCl and EDTA-free complete protease inhibitor cocktail tablets (Roche)) followed by centrifugation at 45,000 r.p.m. for 45 min in a Ti-45 rotor at 4 °C. The low-salt wash was repeated twice, followed by three washes in high-salt buffer (10 mM HEPES pH 7.5, 10 mM $MgCl_2$, 20 mM KCl, 1.0 M NaCl). The membranes were resuspended in 25 mM HEPES pH 7.5, 10 mM $MgCl_2$, 20 mM KCl, 30% (w/v) glycerol, flash-frozen in liquid nitrogen and stored at −80 °C until purification. For purification, membranes were thawed on ice, incubated with 10 mM sodium succinate for 1 h at 4 °C and then solubilized in 50 mM HEPES pH 7.5, 800 mM NaCl, 5 mM sodium succinate, 10% (w/v) glycerol, 1% (w/v) lauryl maltose neopentyl glycol (LMNG) and 0.2% (w/v) cholesteryl hemisuccinate (CHS) at 4 °C for 4 h.

Insoluble material was removed by centrifugation at 45,000 r.p.m. in a Ti-45 rotor for 90 min, imidazole (pH 7.5) was added to a final concentration of 20 mM and the solution was incubated with TALON resin (Clontech) equilibrated in 25 mM HEPES pH 7.5, 800 mM NaCl, 5 mM sodium succinate, 20 mM imidazole, 10% (w/v) glycerol at 4 °C overnight. The resin was washed with 10 column volumes of equilibration buffer (25 mM HEPES pH 7.5, 800 mM NaCl, 1 mM sodium succinate, 20 mM imidazole, 10% (w/v) glycerol, 0.005% (w/v) LMNG, 0.001% (w/v) CHS) followed by 10 column volumes of wash buffer (25 mM HEPES pH 7.5, 800 mM NaCl, 1 mM sodium succinate, 25 mM imidazole, 10% (w/v) glycerol, 0.005% (w/v) LMNG, 0.001% (w/v) CHS). The receptor was eluted with 10 column volumes of elution buffer (25 mM HEPES pH 7.5, 800 mM NaCl, 1 mM sodium succinate, 300 mM imidazole, 10% (w/v) glycerol, 0.005% (w/v) LMNG, 0.001% (w/v) CHS). Protein-containing fractions were pooled and directly used for complex formation with Nanobody6.

Receptor preparations for other assays were further purified by size-exclusion chromatography on a Superdex S200 Increase 10/300GL column (GE Healthcare) equilibrated in 25 mM HEPES pH 7.5, 800 mM NaCl, 10% (w/v) glycerol, 0.005% (w/v) LMNG, 0.001% (w/v) CHS.

Glycerol-free receptor was purified as described above, but glycerol was omitted from the buffer at all steps of membrane preparation and purification.

### K18$^{1.31}$E/K269$^{7.32}$N rat SUCNR1 (humanized rat SUCNR1)

Rat SUCNR1 (residues 2–317) with the two point mutations K18$^{1.31}$E and K269$^{7.32}$N was generated by site-directed mutagenesis using wild-type rat SUCNR1 as a template and cloned into a pIEx/Bac-3 derived plasmid with a haemagglutinin signal sequence followed by a Flag tag at the N terminus of the receptor and a 10×His-tag preceded by a HRV 3C protease cleavage site at the C terminus. Baculovirus generation, expression and purification were performed as described for rat SUCNR1.

### Human BRIL–SUCNR1

The sequence coding for wild-type human SUCNR1 (residues 2–334) was cloned into a pIEx/Bac-3-derived plasmid with a haemagglutinin signal sequence followed by a Flag tag and a cytochrome-B562 (BRIL) at the N terminus of the receptor and a 10×His-tag preceded by a HRV 3C protease cleavage site at the C terminus. Baculovirus generation, expression and purification were performed as described for rat SUCNR1.

### Human SUCNR1

The sequence coding for wild-type human SUCNR1 (residues 2–334) was cloned into a pIEx/Bac-3-derived plasmid with a haemagglutinin signal sequence followed by a Flag tag at the N terminus of the receptor and a 10×His-tag preceded by a HRV 3C protease cleavage site at the C terminus. Baculovirus generation, expression and purification were performed as described for rat SUCNR1.

### Large-scale complex formation of rat SUCNR1 with Nanobody6 for crystallization

Compound 5g (ref. [16]) was added to purified rat SUCNR1 to a final concentration of 1 mM from 100 mM stock solution in DMSO, giving a final DMSO concentration of 1% (v/v). The receptor was mixed with a 1.2 molar excess of purified Nanobody6, incubated on ice for 30 min and concentrated using a 100 kDa molecular weight cut-off concentrator (Millipore) to a final volume of 500 µl. The complex was further purified by size-exclusion chromatography on a S200 Increase 10/300 GL column (GE Healthcare) equilibrated in 25 mM HEPES pH 7.5, 800 mM NaCl, 10% (w/v) glycerol, 0.002% (w/v) LMNG, 0.0004% (w/v) CHS and 25 µM of compound 5g (ref. [16]). Peak fractions were pooled and concentrated using a 100 kDa MWCO concentrator (Millipore) to a final concentration of 40–50 mg $ml^{-1}$. The complex was flash-frozen in liquid nitrogen in small aliquots and stored at −80 °C until crystallization.

### Large-scale complex formation of humanized rat SUCNR1 with Nanobody6 and NF-56-EJ40 for crystallization

NF-56-EJ40 was added to purified humanized rat SUCNR1 to a final concentration of 100 µM from 100 mM stock solution in DMSO, giving a final DMSO concentration of 0.1% (v/v). The receptor was mixed with a 1.1 molar excess of purified Nanobody6, incubated on ice for 30 min and concentrated using a 100 kDa MWCO concentrator (Millipore) to a final volume of 250 µl. The complex was further purified by size-exclusion chromatography on a S200 Increase 10/300 GL column (GE Healthcare) equilibrated in 25 mM HEPES pH 7.5, 800 mM NaCl, 10% (w/v) glycerol, 0.002% (w/v) LMNG, 0.0004% (w/v) CHS and 20 µM of NF-56-EJ40. Peak fractions were pooled and concentrated using a 100 kDa MWCO concentrator (Millipore) to a final concentration of 37 mg ml$^{-1}$. The complex was flash-frozen in liquid nitrogen in small aliquots and stored at −80 °C until crystallization.

### Lipidic cubic phase crystallization

The rat SUCNR1-Nanbody6 complex was reconstituted in lipidic cubic phase (LCP) by mixing protein at 40–50 mg ml$^{-1}$ with monoolein:cholesterol (9:1) (w:w) at a 2:3 ratio (v:w) in 50 µl Hamilton syringes using the two-syringe method[28]. Crystallization trials were performed using Laminex glass sandwich plates (Molecular Dimensions) with a 200-µm spacer and dispensed using a Mosquito LCP robot (Labtech TTP). Protein-laden LCP (50 nl) was covered with 800 nl of precipitant and incubated at 20 °C in an RI-1000 imager (Formulatrix). The first crystals appeared within 24 h and grew to a maximum size of 60 µm × 10 µm × 10 µm in 2 weeks in 100 mM sodium citrate pH 4.8–5.4, 24–30% (w/v) PEG400, 50 mM NaSCN, 2.5% (v/v) 2,5-hexanediol, 1% (v/v) DMSO. Crystals were directly collected from the LCP bolus with MiTeGen micromount loops and flash-frozen in liquid nitrogen.

The humanized rat SUCNR1–Nanbody6–NF-56-EJ40 complex was reconstituted in LCP by mixing protein at 37 mg ml$^{-1}$ with monoolein:cholesterol (9:1) (w:w) at a 2:3 ratio (v:w) in 50 µl Hamilton syringes using the two-syringe method[28]. Crystallization trials were performed using Swissci Xpol glass sandwich plates (Swissci) with a 200-µm spacer and dispensed using a Mosquito LCP robot (Labtech TTP) with active humidification of the pipetting chamber. Protein-laden LCP (50 nl) was covered with 800 nl of precipitant and incubated at 20 °C in an RI-1000 imager (Formulatrix). The first crystals appeared within 12 h and grew to a maximum size of 60 µm × 60 µm × 60 µm in 5 days in 50 mM 2-[(2-amino-2-oxoethyl)-(carboxymethyl)amino]acetic acid (ADA) pH 7.0, 28% (w/v) poly(ethylene glycol) monomethyl ether (PEG MME) 550, 0.55 M $(NH_4)_2SO_4$, 100–400 µM NF-56-EJ40 and 1–4% (v/v) DMSO. Crystals were directly collected from the LCP bolus with MiTeGen micromount loops and flash-frozen in liquid nitrogen.

### Data collection, structure solution and refinement

Data for the rat SUCNR1–Nanobody6 complex were collected at PXI at the Swiss Light Source, Villigen, Switzerland using a 10-µm and a 20-µm diameter beam at a wavelength of 0.999 Å on a Dectris Eiger-16M detector. Crystals were exposed for 0.1 s per 0.1° oscillation per frame using an attenuated beam to reduce radiation damage. Datasets were integrated, scaled and merged using XDS and XSCALE in autoPROC and aP_Scale (Global Phasing). The final dataset was merged from 18 crystals and anisotropic-scaled with diffraction limits of 2.959 Å, 2.088 Å and 2.345 Å. The structure was solved by molecular replacement in Phaser[29] using the structure of the P2Y1 receptor without the rubredoxin fusion as a search model (PDB ID: 4XNV). An initial structural model was built using Autobuild/Resolve in Phenix[30] and further adjusted by repetitive rounds of manual model building in Coot[31] and refinement against the anisotropic-scaled data in Buster (Global Phasing). 99.76% of residues are within the allowed regions of the Ramachandran plot, with 0.24% being outliers and a clash-score of 5. The final model lacks residues 1–5 at the N terminus, residues 160–167 in ECL2 and residues 214–223 in ICL3 of rat SUCNR1 and the first 9 residues at the N terminus of Nanbody6.

Data for the humanized rat SUCNR1–Nanobody6–NF-56-EJ40 complex were collected at PXI at the Swiss Light Source using a 20-µm diameter beam at a wavelength of 1.0002 Å on a Dectris Eiger-16M detector. Crystals were exposed for 0.1 s per 0.2° oscillation per frame using an attenuated beam to reduce radiation damage. Datasets were integrated, scaled and merged using XDS and Aimless in autoPROC and aP_Scale (Global Phasing). The final dataset was obtained from a single crystal and anisotropic-scaled with diffraction limits of 2.327 Å, 1.940 Å and 1.959 Å. The structure was solved by molecular replacement in Phaser[29] using the structure of the wild-type rat SUCNR1–Nanobody6 complex as a search model. The structural model was further adjusted by repetitive rounds of manual model building in Coot[31] and refinement against the anisotropic-scaled data in Buster (Global Phasing). 100% of residues are within the allowed regions of the Ramachandran plot, and the model has a clash-score of 4. The final model lacks residues 1–6 at the N terminus, residues 215–223 in ICL3 and residues 257–261 in ECL3 of humanized rat SUCNR1 and the first 8 residues at the N terminus of Nanobody6.

### Molecular docking

Homology modelling of human SUCNR1 was performed with Prime[32] (Schrödinger, v.2018-3). The coordinates of the high-resolution structure of the full-length wild-type rat isoform of SUCNR1 were used as a template. Missing residues 160–167 in ECL2 in the crystal structure were reconstructed by Prime. Default parameters were used. Compounds were prepared for docking using Corina (MN-AM, v.4.2.0) for initial conformation generation and protonated at pH 7.4 with blabber_sd utility from the MoKa package[33] (Molecular Discovery, v.2.6.6). Dockings were performed with Glide[34,35] (Schrödinger, v.2018-3) with the SP mode and default parameters. Poses were visually inspected for selection.

### Nano-differential scanning fluorimetry thermostability assay

Protein samples at 0.4 mg ml$^{-1}$ were prepared from 4.0 mg ml$^{-1}$ stock by adding assay buffer with or without glycerol (25 mM HEPES pH 7.5, 800 mM NaCl, 0–10% (w/v) glycerol, 0.005% (w/v) LMNG, 0.001% (w/v) CHS) and ligand as required. The ligand added corresponded to no more than 5% of the final assay. Samples were incubated on ice and manually loaded into standard nano-differential scanning fluorimetry (nanoDSF) grade capillaries (NanoTemper Technologies). Experiments were performed with a Prometheus nano-DSF instrument (NanoTemper Technologies) with a temperature gradient from 20 °C to 95 °C and a temperature slope of 2.0 °C or 2.5 °C min$^{-1}$. Data were processed and analysed using the PR.ThermControl Software (NanoTemper Technologies).

For the assessment of compound stability, nano-DSF was performed at a protein concentration of 0.2 mg ml$^{-1}$ in 25 mM HEPES pH 7.5, 800 mM NaCl, 10% (w/v) glycerol, 0.01% (w/v) LMNG, 0.002% (w/v) CHS in the presence of 50 µM compound and 0.5% (v/v) DMSO.

### Analytical size-exclusion chromatography

Samples were analysed on an HPLC 1100 instrument (Agilent) equipped with a Zenix-C SEC-300 column (Sepax Technologies) and a photodiode detector (Wyatt) in 50 mM MES pH 6.0, 500 mM NaCl, 0.01% (w/v) LMNG with a flow rate of 0.3 ml min$^{-1}$ or on an AKTA Micro equipped with a S200 Increase 3.2/300 column (GE Healthcare) in 25 mM HEPES pH 7.5, 800 mM NaCl, 10% (w/v) glycerol, 0.01% (w/v) LMNG and 0.002% (w/v) CHS at a flow rate of 50 µl min$^{-1}$.

### [$^{35}$S]GTPγS assay

Membranes were prepared from stably transfected rat SUCNR1 CHO-K1 and human S1P1 CHO-K1 cell lines (Novartis) and from Chem1 cells expressing human SUCNR1 (Millipore). Cell cultures were rinsed with PBS and the cells were scraped off the flasks in ice-cold 20 mM HEPES buffer pH 7.4, 10 mM EDTA containing protease inhibitors (Roche). The pellet obtained after centrifugation for 30 min at 17,500*g* was resuspended in

ice-cold buffer as above supplemented with 100 mM NaCl and homogenized using a PT 1300 D homogenizer (Polytron) at 25,000 r.p.m. for three intervals of 20 s each. The homogenate was centrifuged at 39,000*g* for 40 min at 4 °C, the pellet resuspended and homogenized again at 25,000 r.p.m. for 20 s (PT 1300 D homogenizer). Aliquots were stored at −80 °C.

For [$^{35}$S]GTPγS assays, membranes were resuspended in assay buffer (20 mM HEPES pH 7.4, 100 mM NaCl, 10 mM $MgCl_2$, 25 μg ml$^{-1}$ saponin, 100 μM GDP, 0.1% fat-free bovine serum albumin (BSA)) containing wheat germ agglutinin (WGA)-coated scintillation proximity assay (SPA) beads (Perkin Elmer). The final assay mixture in 96-well Optiplates (215 μl final volume; Perkin Elmer) contained 15 μg rat or human SUCNR1 or 3 μg human S1P1R membranes, 0.2 mg (rat or human SUCNR1) or 1 mg (human S1P1R) WGA SPA beads, 200 pM [$^{35}$S]GTPγS (Perkin Elmer), 20 mM HEPES pH 7.4, 100 mM NaCl, 10 mM $MgCl_2$, 25 μg ml$^{-1}$ saponin, 10 μM GDP, 0.1% fat-free BSA, receptor agonist (succinate or S1P) and test agents as indicated (Nanobody6 or glycerol). Assay plates were sealed and incubated with continuous shaking for 60 min at room temperature. Afterwards the plate was centrifuged for 10 min at 1,200 r.p.m. and the radioactivity counted using a TopCount NXT instrument (Perkin Elmer).

### [$^3$H] Radioligand-binding assay

Cytomegalovirus promoter-based SUCNR1 expression plasmids containing a haemagglutinin signal sequence followed by a Flag tag at the N terminus of the receptor and a 10×His-tag preceded by a HRV 3C protease cleavage site at the C terminus were transiently transfected (Lipofectamine 2000; Life Technologies) into HEK293FT cells (Life Technologies). Membranes were prepared two days after transfection as described above and homogenized in Krebs-Tris buffer (20 mM Tris-HCl pH 7.4, 118 mM NaCl, 5.6 mM glucose, 1.2 mM $KH_2PO_4$, 4.7 mM KCl, 1.2 mM $MgSO_4$, 1.8 mM $CaCl_2$). Assay mixtures (200 μl) in 96-well multiscreen filter plates MSFBN6B10 (Millipore) contained 20 μg cell membranes, 50 nM [$^3$H]NF-56-EJ40 (963 GBq mmol$^{-1}$; synthesized at Novartis) and test compounds and were incubated for 1 h at room temperature. After two washes with 200 μl ice-cold Krebs-Tris buffer the plates were air-dried, the bottom of the plates sealed and 40 μl of scintillation fluid was added (MicroScint PS, Perkin Elmer). Radioactivity was counted using a Topcount NXT (Perkin Elmer). Half-maximum inhibitory concentration ($IC_{50}$) values were determined from eight-point concentration–response curves and $K_i$ values were calculated using the Cheng–Prusoff equation.

### Figure preparation

Protein structure figures were prepared using PyMOL incentive 2.0.7. (Schrodinger).

Sequences were aligned with ClustalOmega[36] and figures were prepared using ESPript3.0[37]. ConSurf[38] was used to visualize sequence conservation on structures. [$^{35}$S]GTPγS assay data were analysed and figures prepared using GraphPad Prism.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

Structure factors and coordinates of the rat SUCNR1–Nanobody6 and the SUCNR1(K18$^{1.31}$E/K269$^{7.32}$N)–Nanobody6–NF-56-EJ40 complex structures have been deposited in the Protein Data Bank (PDB) under accession codes 6IBB and 6RNK, respectively. All source data associated with the paper (in addition to those deposited) are provided as Supplementary Information.


28. Caffrey, M. & Cherezov, V. Crystallizing membrane proteins using lipidic mesophases. *Nat. Protoc.* **4**, 706–731 (2009).
29. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
30. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
31. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
32. Jacobson, M. P. et al. A hierarchical approach to all-atom protein loop prediction. *Proteins* **55**, 351–367 (2004).
33. Milletti, F., Storchi, L., Sforna, G. & Cruciani, G. New and original p$K_a$ prediction method using grid molecular interaction fields. *J. Chem. Inf. Model.* **47**, 2172–2181 (2007).
34. Halgren, T. A. et al. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **47**, 1750–1759 (2004).
35. Friesner, R. A. et al. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **47**, 1739–1749 (2004).
36. McWilliam, H. et al. Analysis tool web services from the EMBL-EBI. *Nucleic Acids Res.* **41**, W597–W600 (2013).
37. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–W324 (2014).
38. Ashkenazy, H. et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44**, W344–W350 (2016).

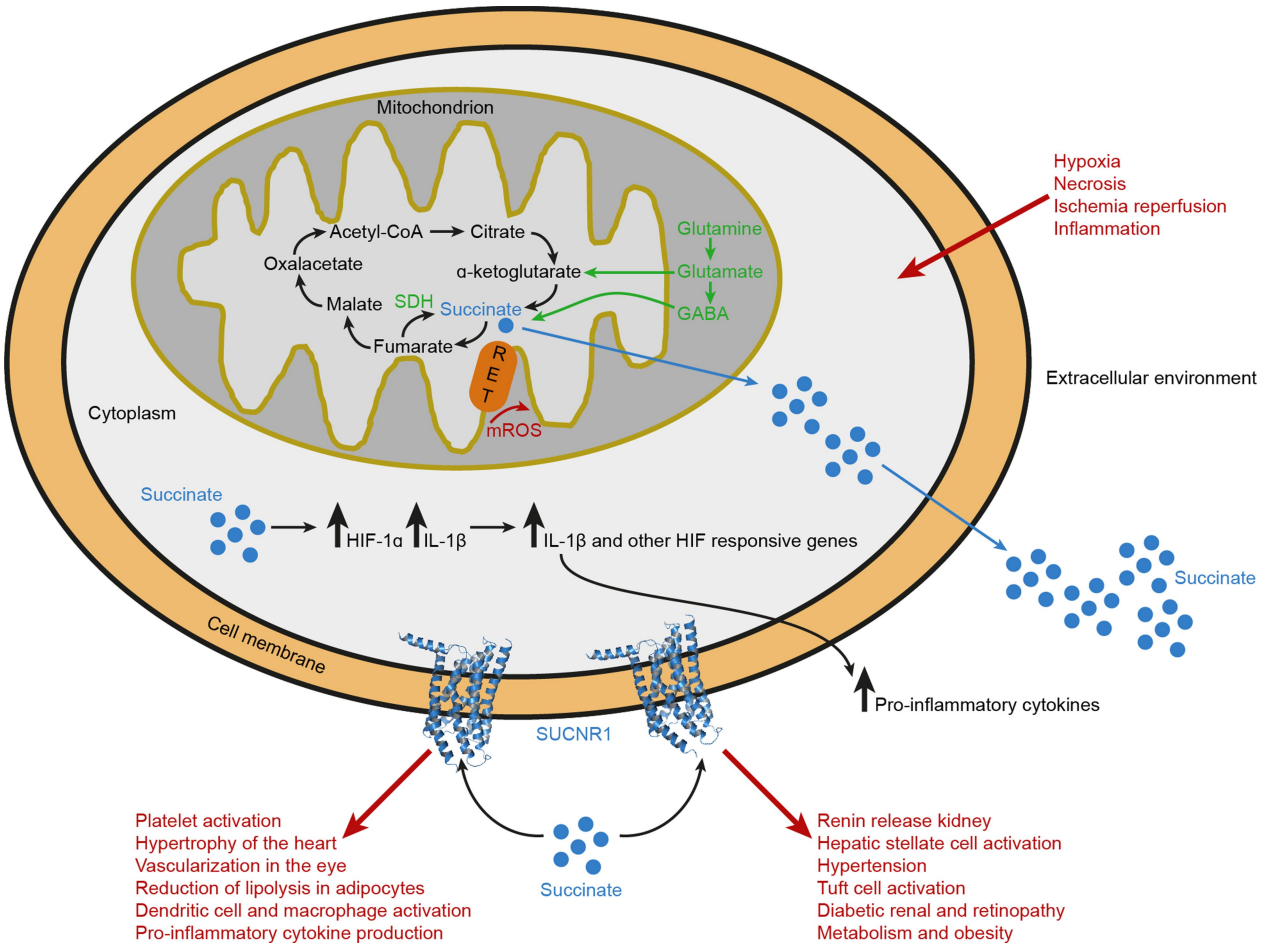


**Acknowledgements** We thank C. Schleberger for help with implementing crystallographic data processing tools; T. Huber, R. Link and A. Winterhalter for help with nanobody generation; P. Loesle for biochemical assay support; S. Haenni and S. Holzinger for molecular biology support; and E. Loetscher for providing the rat SUCNR1-CHO cell line.

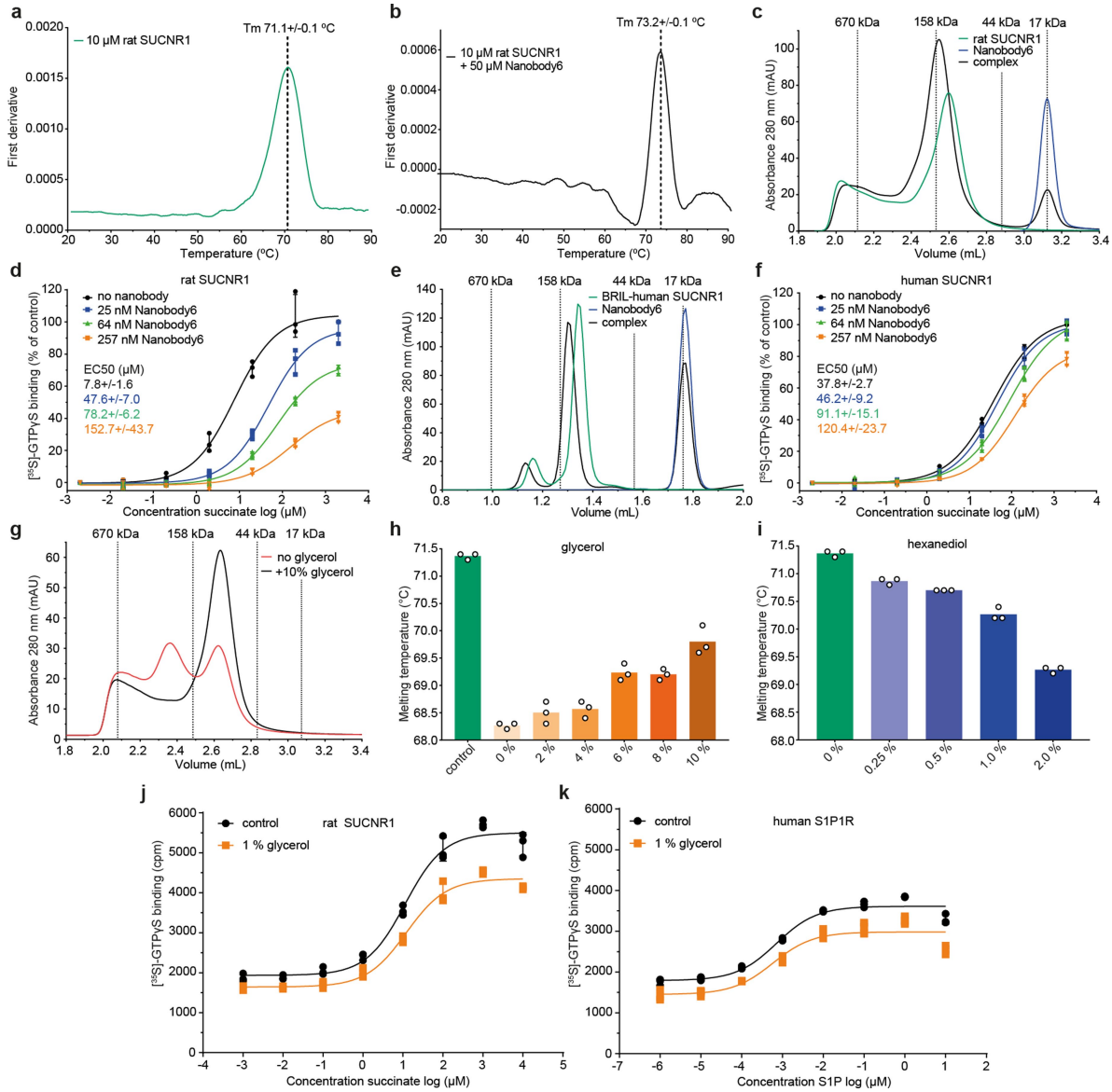

**Author contributions** M.H., A.L.-E., K.K., M.D. and V.P.J. designed the experiments. G.R., J.B. and V.P.J. performed initial protein expression and purification screening and established receptor purification protocols. M.H. expressed and purified all proteins, established complex purification and crystallization, solved and analysed structures, performed thermostability assays and proposed structure-based mutagenesis. D.F. and K.K. established and performed radioligand-binding and GTPγS assays. M.D. initiated and supervised nanobody generation. S.C. designed compounds and built human SUCNR1 homology models. F.S. built structural models of the humanized rat and human SUCNR1 and performed docking analysis. N.G. designed and synthesized compounds. F.F. designed expression constructs. M.H., K.K., A.L.-E. and V.P.J. wrote the manuscript with input from all authors.

**Competing interests** All authors are employees of Novartis Pharma AG.

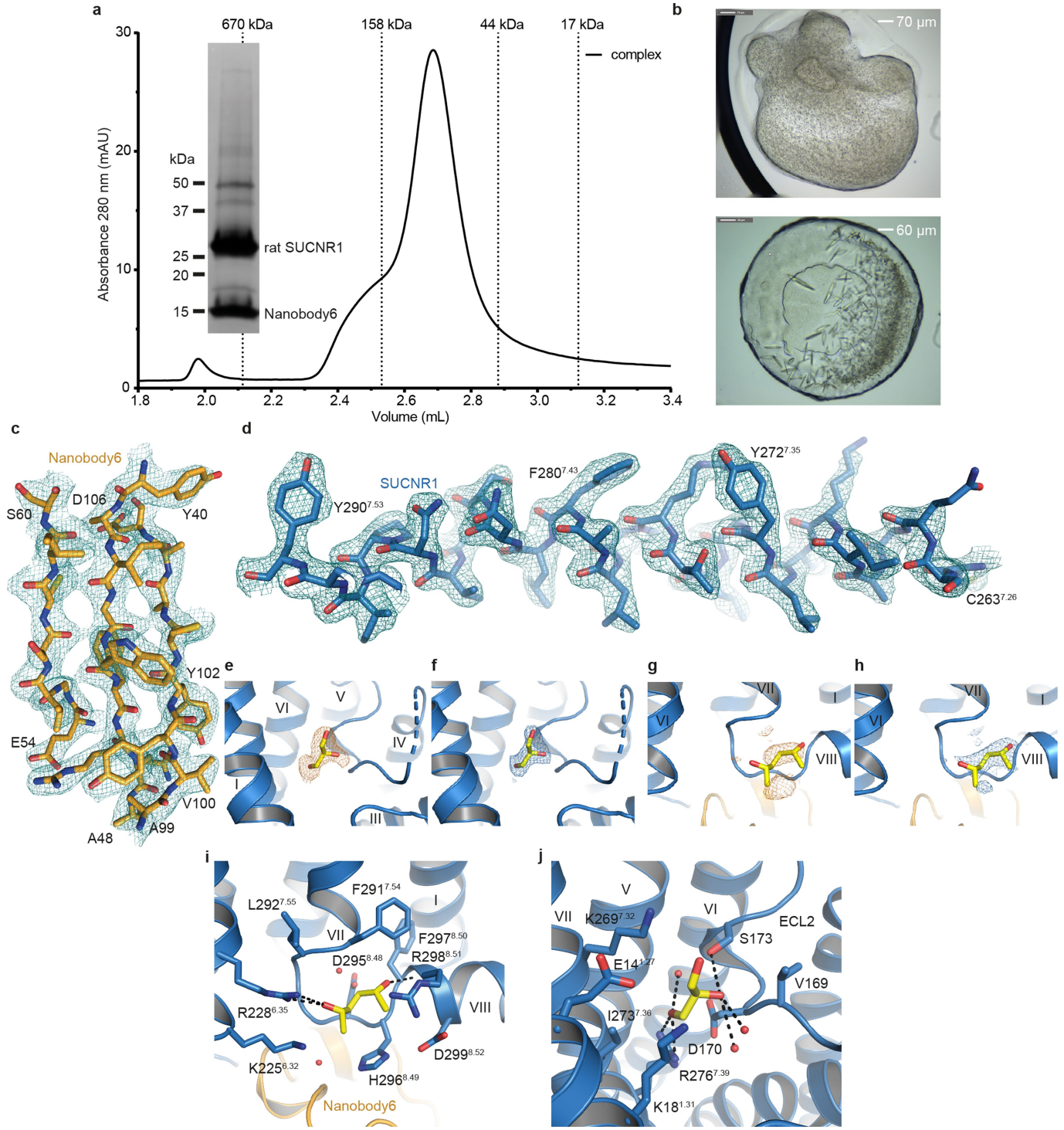

**Additional information**
**Supplementary information** is available for this paper at https://doi.org/10.1038/s41586-019-1663-8.
**Correspondence and requests for materials** should be addressed to M.H., K.K. or V.-P.J.
**Peer review information** *Nature* thanks Edward Chouchani, Mike Murphy, Janos Peti-Peterdi and Gebhard Schertler for their contribution to the peer review of this work.
**Reprints and permissions information** is available at http://www.nature.com/reprints.



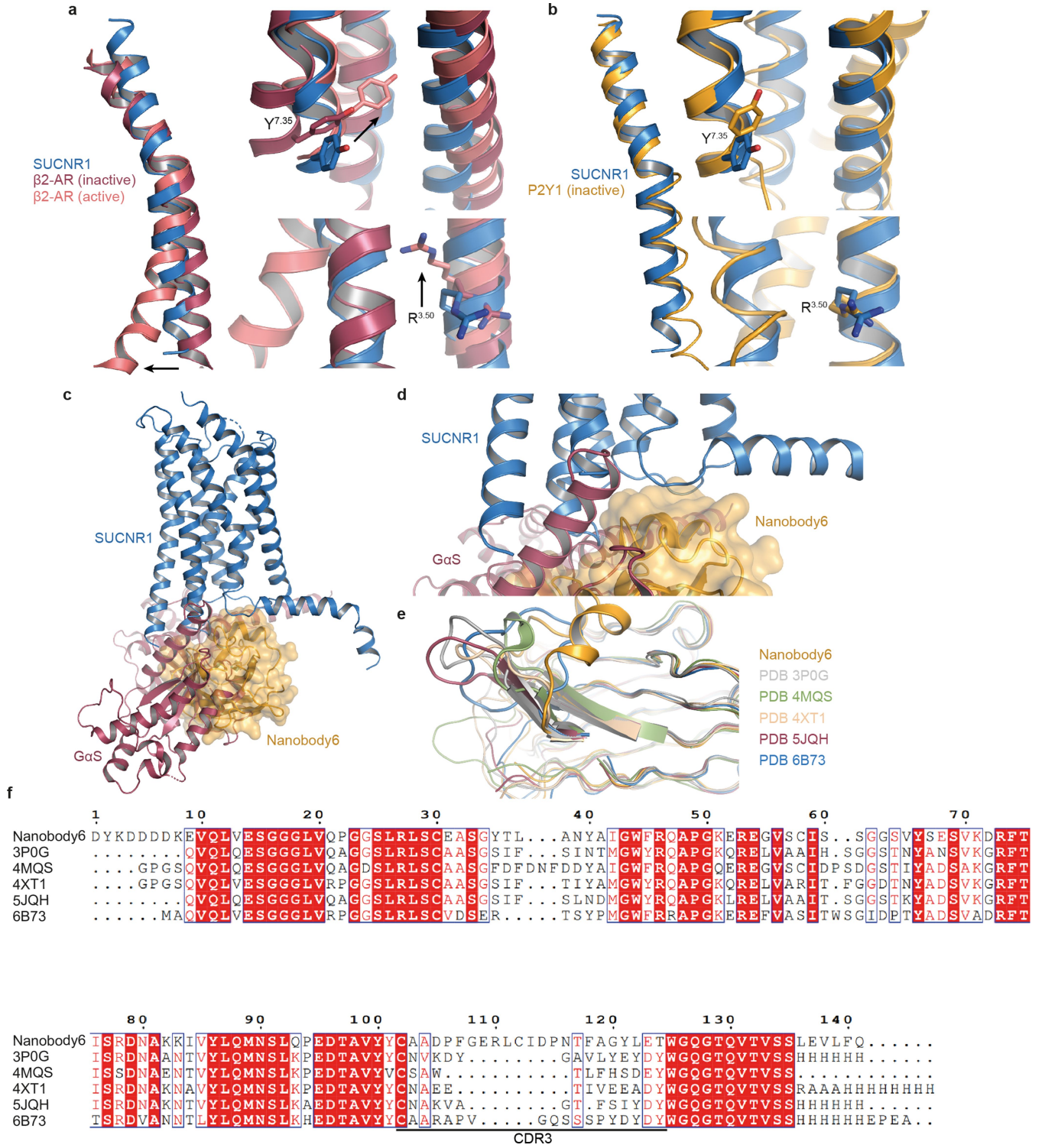

**Extended Data Fig. 1 | The succinate–SUCNR1 signalling axis.** Levels of the Krebs cycle intermediate succinate are increased under certain conditions such as hypoxia, necrosis, ischaemia reperfusion and inflammation. The ways in which succinate concentrations increase in the mitochondrion are shown in green. Mitochondrial reactive oxygen species result from the reversed electron transport (RET) chain driven by an increase in succinate. Succinate is transported into the cytoplasm, where it can stabilize HIF1α and increase the expression of genes that have HIF-responsive elements, such as *IL1B*. Further succinate is exported into the local extracellular environment, where it accumulates and binds and activates SUCNR1. Several of the consequences of this are shown in red. GABA, γ-aminobutyric acid; HIF, hypoxic inducible factor; IL-1, interleukin-1; mROS, mitochondrial reactive oxygen species; SDH, succinate dehydrogenase.




**Extended Data Fig. 2 | Interaction of Nanobody6 with rat and human SUCNR1 and characterization of the effects of glycerol and 2,5-hexanediol on rat SUCNR1. a**, **b**, Nanobody6 increases the thermal stability of rat SUCNR1 in nano-DSF thermal shift assays. The average melting curves in the absence (**a**) and presence (**b**) of Nanobody6 are shown ($n = 3$; technical replicates). The experiment was repeated independently 3 times with similar results. **c**, Analytical size-exclusion chromatography shows a clear shift in the peak of the rat SUCNR1–Nanobody6 complex compared to the peak of the receptor alone. The complex samples contain a 1.2 molar excess of Nanobody6 over receptor. One of $n = 2$ independent experiments is shown. **d**, [$^{35}$S]GTPγS assay on wild-type rat SUCNR1 in the absence or the presence of increasing concentrations of Nanobody6. The average curves of $n = 3$ independent experiments are shown; data are mean ± s.d. Average half-maximum effective concentration ($EC_{50}$) values from $n = 3$ independent experiments are listed; data are mean ± s.d. **e**, Analytical size-exclusion chromatography shows a clear shift in the peak of the N-terminal BRIL-fused human SUCNR1–Nanobody6 complex compared to the peak of the receptor alone. One of $n = 2$ independent experiments is shown. **f**, [$^{35}$S]GTPγS assay on wild-type human SUCNR1 in the absence or the presence of increasing concentrations of Nanobody6. The average curves of $n = 3$ independent experiments are shown and average $EC_{50}$ values from $n = 3$ independent experiments are listed; data are mean ± s.d. **g**, Analytical size-exclusion chromatography of rat SUCNR1 purified in the absence or the presence of 10% glycerol. One of $n = 2$ independent experiments is shown. **h**, Glycerol increases the thermal stability of rat SUCNR1 as evidenced from nano-DSF assays ($n = 3$ technical replicates; bars represent mean values; individual data points are indicated by circles). The control sample was purified in the presence of 10% glycerol. All other samples contain rat SUCNR1 purified without glycerol, to which the respective final glycerol concentration was added. The experiment was repeated independently twice with similar results. **i**, 2,5-Hexanediol decreases the thermal stability of rat SUCNR1 in nano-DSF assays ($n = 3$ technical replicates; bars represent mean values; individual data points are indicated by circles). The experiment was repeated independently twice with similar results. **j**, **k**, [$^{35}$S]GTPγS assay of rat SUCNR1 (**j**) and human S1P1R (**k**) in the absence or the presence of 1% (109 mM) glycerol. The average curve of $n = 3$ independent experiments and the individual data points of each experiment are shown. Data are mean ± s.d.

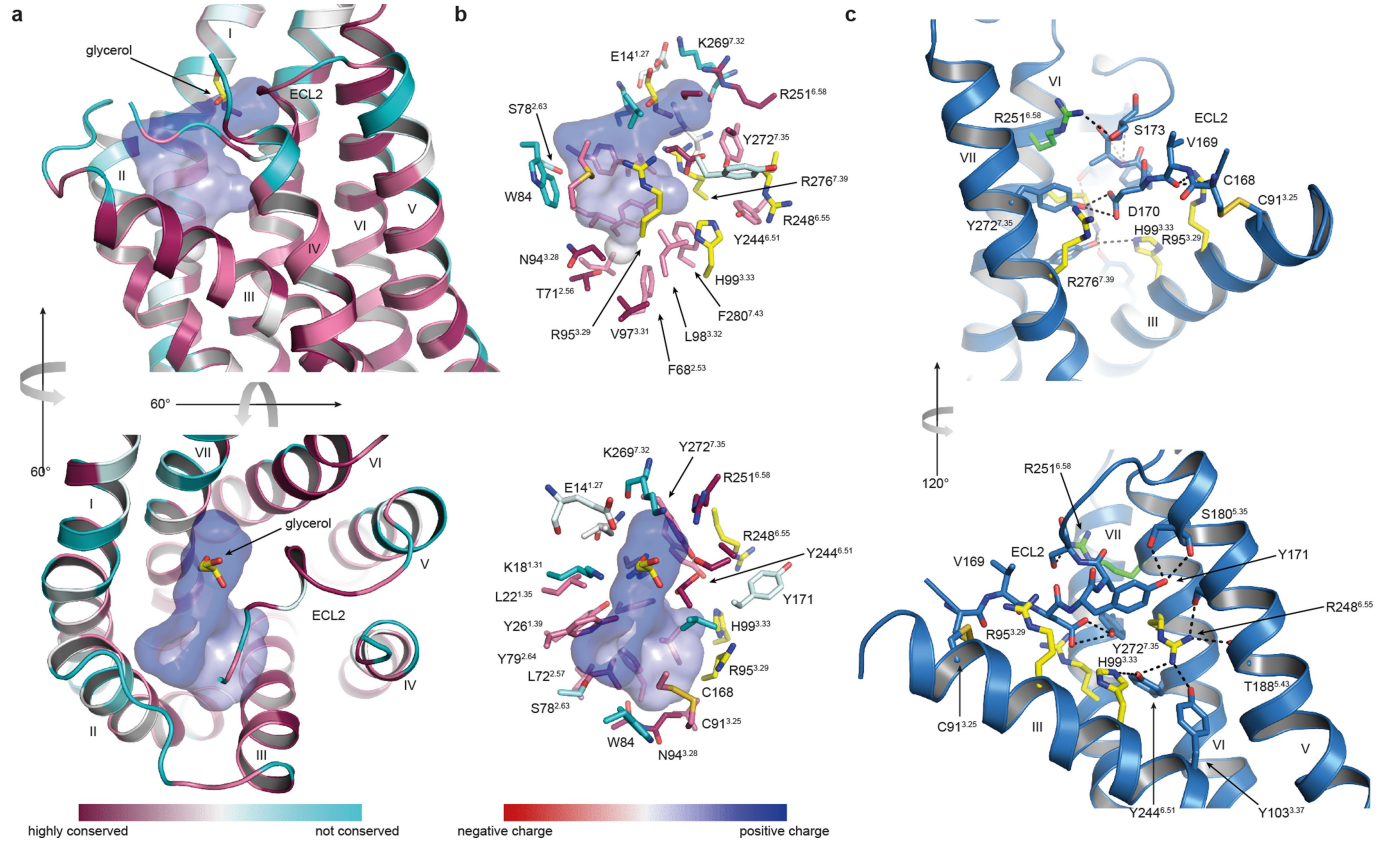



**Extended Data Fig. 3 | Purification, crystallization and electron-density map quality of the rat SUCNR1–Nanobody6 complex, with detailed binding modes of glycerol and 2,5-hexanediol. a**, Analytical size-exclusion chromatography and SDS–PAGE analysis of crystallization samples of the rat SUCNR1–Nanobody6 complex. Shown is a representative experiment of $n = 5$ independent experiments with similar results. For gel source data, see Supplementary Fig. 1a. **b**, Initial crystallization hits for the rat SUCNR1–Nanobody6 complex (top) and optimized crystals used for data collection (bottom). Shown are representative experiments of $n = 20$ independent experiments. **c**, The $2F_o - F_c$ electron density map contoured at $1.5\sigma$ for a part of Nanobody6. **d**, The $2F_o - F_c$ electron-density map contoured at $1.5\sigma$ for helix VII in rat SUCNR1. **e**, **f**, $F_o - F_c$ composite omit map for glycerol (**e**) and corresponding $2F_o - F_c$ electron-density map after refinement (**f**). Both maps are contoured at $1.5\sigma$. **g**, **h**, $F_o - F_c$ composite omit map for 2,5-hexanediol (**g**) and corresponding $2F_o - F_c$ electron-density map after refinement (**h**). Both maps are contoured at $1.5\sigma$. **i**, **j**, Detailed views of the side-chain environment around 2,5-hexanediol (**i**) and glycerol (**j**). Hydrogen bonds are indicated by dashed lines. For clarity, only residues within a distance of 4 Å are shown.

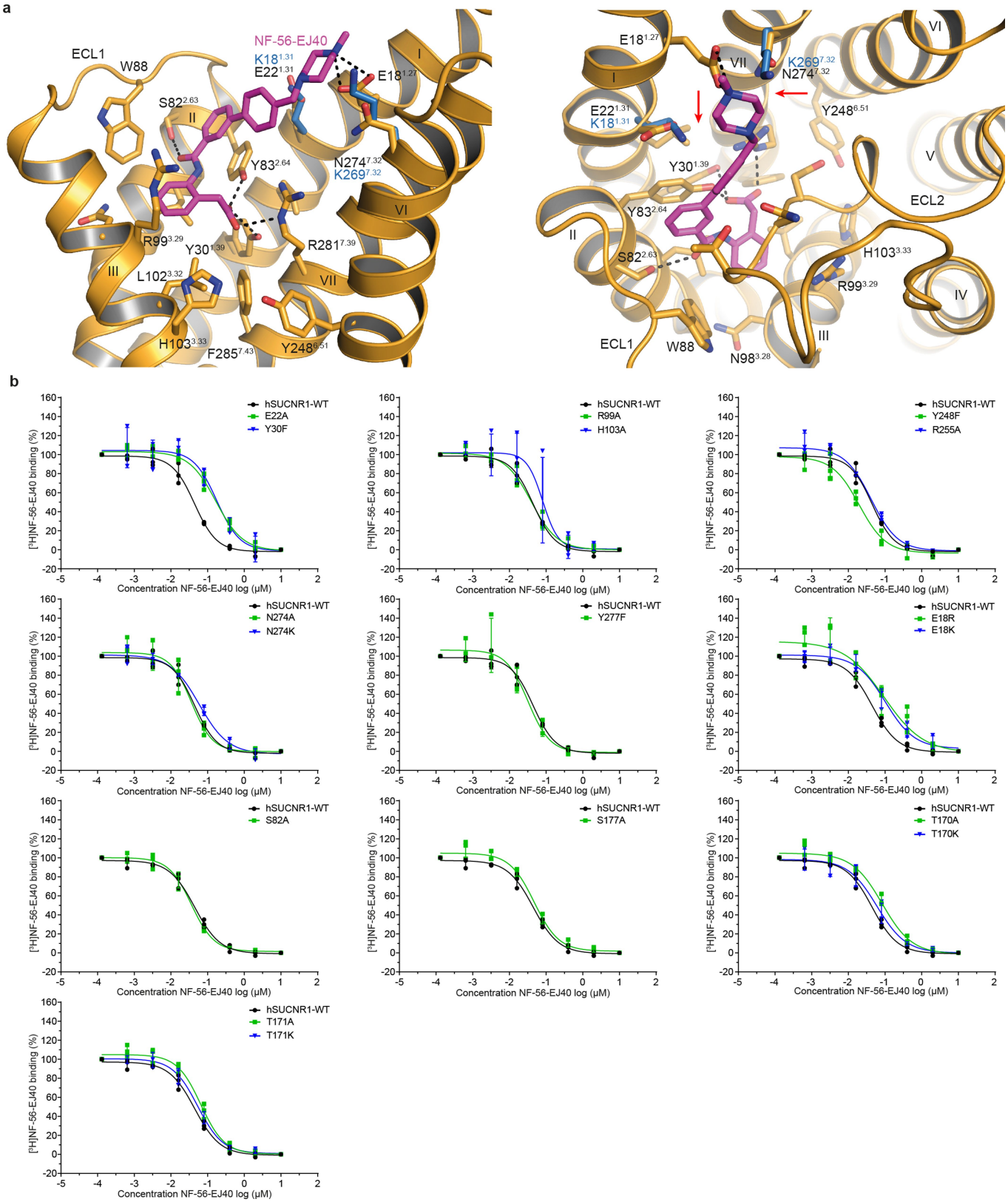



**Extended Data Fig. 4 | Rat SUCNR1 adopts an inactive conformation in complex with Nanobody6, which binds to the intracellular side via an extended CDR3 loop. a**, Structural alignment of helix VI in rat SUCNR1 and the active and inactive states of β2-AR, with the positions of key residues as hallmarks for inactive and active receptor states. SUCNR1 is shown in blue, inactive β2-AR in red and active β2-AR in pink. Alignment of helix VI (left) and side-chain positions of key residues ($R^{3.50}$ and $Y^{7.35}$) (right) indicate an inactive state for rat SUCNR1. **b**, Structural alignment of helix VI and key residues ($R^{3.50}$ and $Y^{7.35}$) of rat SUCNR1 (blue) and the P2Y1 receptor (orange) in the inactive conformation. **c**, Superposition of the rat SUCNR1–Nanobody6 complex with the $G\alpha_s$ subunit from the β2-AR $G_s$-protein trimer structure (PDB ID: 3SN6). Rat SUCNR1 is shown in blue, Nanobody6 in orange and the $G\alpha_s$ subunit in red. The G-protein and the Nanobody6-binding site partially overlap. **d**, Magnified view of the overlap between the G-protein and the Nanobody6-binding site of rat SUCNR1. **e**, Structural alignment of nanobodies used to crystallize GPCRs. The extended CDR3 of Nanobody6 forms a helical secondary structure. **f**, Sequence alignment of the GPCR-stabilizing nanobodies shown in **e**. The PDB IDs of the respective GPCR–nanobody complex structures are listed and the CDR3 region is highlighted by a black bar.

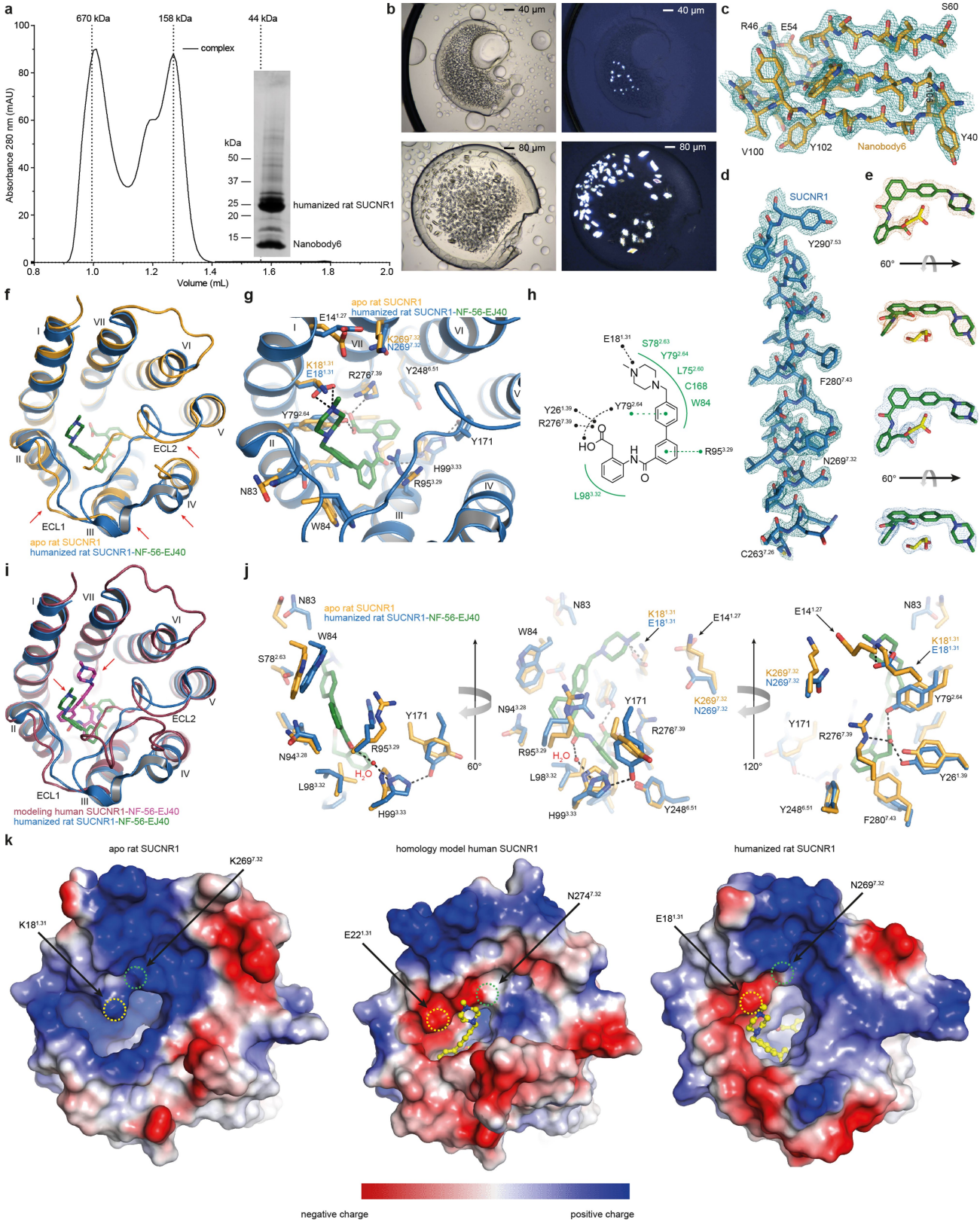

**Extended Data Fig. 5 | Sequence alignment of rat, human and mouse SUCNR1 and ConSurf analysis. a**, The helical structure elements as observed in the crystal structure of apo rat SUCNR1 are indicated. Sequences corresponding to ECL1 and ECL2 are boxed in blue and the non-conserved residues K18$^{1.31}$ and K269$^{7.32}$ are marked by blue arrows. Yellow arrowheads indicate residues that were previously reported to be involved in succinate-induced receptor activation. Green dots indicate residues that are involved in NF-56-EJ40 binding in the humanized rat SUCNR1 structure. **b**, The sequence alignment of SUCNR1 from various species is colour-coded from turquoise (variable) to dark pink (conserved), similar to the colours used in Extended Data Fig. 6a, b, on the basis of analysis with ConSurf. Residues involved in the binding of Nanobody6 are indicated by blue triangles.




**Extended Data Fig. 6 | Potential ligand-binding site in apo SUCNR1 is partially occluded by ECL2. a**, Side view (top) and top view (bottom) of the hydrophobic pocket located below the glycerol molecule. The sequence conservation within the SUCNR1 receptor family is indicated by colour, ranging from turquoise for highly variable residues to dark pink for highly conserved residues. The hydrophobic pocket is shown as a surface, colour-coded by charge. The glycerol molecule is shown as yellow sticks. **b**, The same orientations as in **a** are shown. Residues forming the deep hydrophobic pocket are shown as sticks, colour-coded by sequence conservation as in **a**. Residues that were previously reported to be involved in succinate-induced receptor activation are coloured yellow. **c**, Residues in the environment of ECL2 are shown as sticks and hydrogen bonds are shown as dashed black lines. Residues that have previously been reported to have an effect on succinate binding by the receptor are shown in yellow[2]; R251$^{6.58}$, which was identified in a second study[15], is shown in green.




**Extended Data Fig. 7 | Identification of critical residues that impart species selectivity for antagonist binding. a**, Side view (left) and top view (right) of the potential binding mode of the antagonist NF-56-EJ40 (shown in pink), obtained by molecular modelling based on the crystal structure of apo rat SUCNR1. Residues within 4 Å of NF-56-EJ40 are shown as sticks. Red arrows point towards two sites in which potential steric clashes may occur with rat-SUCNR1-specific residues K18$^{1.31}$ and K269$^{7.32}$, which are shown in blue. For clarity, ECL2, helix IV and helix V are omitted in the side view. **b**, Radioligand competition binding experiments with unlabelled NF-56-EJ40 on human SUCNR1 mutant proteins. Curves were calculated from $n = 3$ independent experiments; data are mean ± s.d. Individual data points are shown.




**Extended Data Fig. 8** | See next page for caption.

**Extended Data Fig. 8 | Purification, crystallization and electron-density map quality of the rat SUCNR1–Nanobody6–NF-56-EJ40 complex and structural changes in SUCNR1 after antagonist binding. a**, Analytical size-exclusion chromatography and SDS–PAGE analysis of crystallization samples of the humanized rat SUCNR1–Nanobody6–NF-56-EJ40 complex. Shown is a typical result from $n = 2$ independent experiments. Although partial complex aggregation was observed, this did not interfere with crystallization. For gel source data, see Supplementary Fig. 1b. **b**, Top, initial crystallization hits for the humanized rat SUCNR1–Nanobody6–NF-56-EJ40 complex, shown in normal (left) and cross-polarization (right) imaging modes. Bottom, optimized crystals used for data collection shown in normal (left) or cross-polarization (right) imaging modes. A typical result from $n = 3$ independent experiments is shown. **c**, The $2F_o - F_c$ electron density map contoured at $1.5\sigma$ for a part of Nanobody6. **d**, The $2F_o - F_c$ electron density map contoured at $1.5\sigma$ for helix VII in humanized rat SUCNR1. **e**, $F_o - F_c$ composite omit map (top) for NF-56-EJ40 and glycerol contoured at $1.5\sigma$ is shown in orange. The $2F_o - F_c$ map (bottom) for NF-56-EJ40 and glycerol after refinement contoured at $1.5\sigma$ is shown in blue. **f**, Top view of apo rat SUCNR1 (shown in orange) and humanized rat SUCNR1 (shown in blue) in complex with NF-56-EJ40 (shown as green sticks). Large structural rearrangements are indicated by red arrows. Note that ECL2 is completely structured in the humanized rat SUCNR1 structure. **g**, Top view of the NF-56-EJ40-binding site in humanized rat SUCNR1 overlaid with apo rat SUCNR1. Important side chains around NF-56-EJ40 (shown in green) are shown as sticks and are coloured blue for humanized rat SUCNR1 or orange for apo wild-type rat SUCNR1. For clarity, only the backbone of the humanized rat SUCNR1 is shown in cartoon representation. **h**, Side chains that directly interact with NF-56-EJ40 via hydrogen bonding, π–π stacking and cation–π stacking are listed in black. The hydrogen-bonding interactions are shown by black dashed lines and the π–π and cation–π interactions are shown by green dashed lines. Additional residues with van der Waals interactions are listed in green, and their interaction surfaces are indicated by solid green lines. **i**, Top view of the humanized rat SUCNR1 (blue) in complex with NF-56-EJ40 (green sticks) and of the apo rat SUCNR1-derived model of human SUCNR1 (red) with the binding mode of NF-56-EJ40 (pink) from molecular-docking studies, for which details are shown in Extended Data Fig. 7a. Note how both NF-56-EJ40 poses differ considerably, as indicated by red arrows. **j**, Detailed views of the NF-56-EJ40-binding site. For clarity, only side chains are shown. Humanized rat SUCNR1 is shown in blue; wild-type rat SUCNR1 is shown in orange. Note the side-chain flips for $R95^{3.29}$, $L98^{3.32}$, $H99^{3.33}$ and Y171 between both structures. **k**, Top view of wild-type rat SUCNR1 structure (left), the apo rat SUCNR1-based homology model of human SUCNR1 in complex with NF-56-EJ40 (middle) and the humanized rat SUCNR1 structure in complex with NF-56-EJ40 (right). The surface is shown coloured by electrostatic charge. The two key positions ($K/E^{1.31}$ and $K/N^{7.32}$) are highlighted by arrows. NF-56-EJ40 is shown in yellow as a ball-and-stick model. Note the differences between the NF-56-EJ40-binding mode determined in the modelled structure and in the crystal structure, and between the surface charge distributions in rat, human and humanized rat SUCNR1.

**Extended Data Table 1 | Data collection and refinement statistics**

| | rat SUCNR1-Nanobody6 (PDB 6IBB)† | humanized rat SUCNR1-NF-56-EJ40 (PDB 6RNK)†† |
|---|---|---|
| **Data collection** | | |
| Space group | $P2_1$ | C2 |
| Cell dimensions | | |
| *a*, *b*, *c* (Å) | 60.22, 164, 63.42 | 76.83, 150.93, 68.23 |
| α, β, γ (°) | 90, 102.634, 90 | 90, 112.42, 90 |
| Resolution (Å)* | 82.00-2.103 (2.288-2.103) | 75.47-1.949 (2.085-1.949) |
| $R_{merge}$* | 0.302 (4.965) | 0.105 (1.579) |
| $I / \sigma I$* | 12.3 (1.4) | 10.4 (1.3) |
| Completeness* (%) | | |
| spherical | 59.4 (7.1) | 77.6 (21.3) |
| ellipsoidal | 92.3 (64.3) | 92.7 (57.0) |
| Redundancy* | 30.9 (33.5) | 7.0 (7.4) |
| **Refinement** | | |
| Resolution (Å) | 41.55-2.12 | 75.47-1.94 |
| No. reflections | 41256 | 40585 |
| $R_{work}$ / $R_{free}$ | 17.8 / 21.6 | 17.7 / 19.9 |
| No. atoms | | |
| Protein | 6643 | 3281 |
| Ligand/ion | 374 | 286 |
| Water | 299 | 242 |
| *B*-factors | | |
| Protein | 56.38 | 42.13 |
| Ligand/ion | 77.95 | 69.17 |
| Water | 61.65 | 61.90 |
| R.m.s. deviations | | |
| Bond lengths (Å) | 0.014 | 0.014 |
| Bond angles (°) | 1.70 | 1.56 |

*Values in parentheses are for the highest-resolution shell.
†Data merged from 18 crystals.
††Data obtained from a single crystal.

## Extended Data Table 2 | Data for the binding of radioligand [$^3$H]NF-56-EJ40 with human and rat SUCNR1 mutants

| mutant | $K_i$ (nM) | $pK_i$ ± s. d. |
|---|---|---|
| Human SUCNR1 | 17.4 | 7.76 ± 0.06 |
| E18A | n. a. | n. a. |
| E18R | 41.6 | 7.43 ± 0.26 |
| E18K | 42.2 | 7.39 ± 0.15 |
| E22A | 67.5 | 7.18 ±0.10 |
| E22K | n. a. | n. a. |
| Y30F | 74.9 | 7.14 ± 0.12 |
| S82A | 14.2 | 7.84 ± 0.06 |
| Y83F | n. a. | n. a. |
| Y83A | n. a. | n. a. |
| C95S | n. a. | n. a. |
| R99A | 16.1 | 7.81 ± 0.13 |
| H103A | 37.3 | 7.54 ± 0.38 |
| T170A | 33.8 | 7.48 ± 0.08 |
| T170K | 23.9 | 7.64 ± 0.13 |
| T171A | 26.4 | 7.58 ± 0.06 |
| T171K | 21.4 | 7.68 ±0.10 |
| C172S | n. a. | n. a. |
| S177A | 19.0 | 7.72 ± 0.03 |
| Y248F | 7.8 | 8.12 ± 0.11 |
| Y248A | n. a. | n. a. |
| R252A | n. a. | n. a. |
| R255A | 16.9 | 7.78 ± 0.06 |
| N274A | 14.4 | 7.86 ± 0.17 |
| N274K | 23.7 | 7.63 ± 0.05 |
| Y277F | 12.7 | 7.90 ± 0.1 |
| Y277A | n. a. | n. a. |
| R281A | n. a. | n. a. |
| E22K, K274N | n. a. | n. a. |
| Rat SUCNR1 | | |
| K18E, K269N | 33.5 | 7.48 ± 0.09 |

n.a., binding signal too low (<300 counts per minute) to determine $K_i$. Values were determined from $n = 3$ independent experiments.

nature research

Corresponding author(s): Veli-Pekka Jaakola

Last updated by author(s): Aug 6, 2019

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☒ | ☐ | A description of all covariates tested |
| ☐ | ☒ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☒ | ☐ | For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☒ | ☐ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒ | ☐ | Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |

*Our web collection on statistics for biologists contains articles on many of the points above.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | Synchrotron data collection software interface at beamline X06SA (PXI, SLS/PSI, Villigen, Switzerland), thermostability data were collected using NanoTemper PR.ThermControl version 2.1. Size exclusion chromatography data were collected using UNICORN version 6.3.2. and ChemStation Rev.B.04.03-SP1 on Agilent HPCL systems. |
| Data analysis | X-ray diffraction data were analyzed with XDS and XSCALE Version May 1, 2016 (BUILT 20160617) and AUTOPROC utilizing Pointless version 1.11.3, AIMLESS version 0.5.32, CCP4 version 7.0.0.44 and Staraniso version 1.9.6 (20170920). The structure was solved by MR in Phaser Version 2.8.0 as implemented in Phenix Version 1.15.2-3472 and the structure was refined in Buster Version 2.11.07. An initial structural model was built using AUTOBUILD/RESOLVE in Phenix Version 1.15.2-3472 and further model building was performed in COOT version 0.8.9.1. Structural figures were prepared with PyMol incentive 2.0.7. Biochemical assay data were analyzed in GraphPad Prism version 7.04 and Microsoft Excel 2016. Thermostability assay data were analyzed with NanoTemper PR.ThermControl Software version 2.1 and GraphPad Prism version 7.04 and GraphPad Prism version 8.1.2. For homology modeling and molecular docking, Prime (version 2018-3), Corina (version 4.2.0), blabber_sd from the MoKa package (version 2.6.6.) and Glide (version 2018-3) were used. Sequences were aligned with ClustalOmega (https://www.ebi.ac.uk/Tools/msa/clustalo/) and sequence alignment figures prepared using ESPript3.0. |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

X-ray structure coordinates and structure factors have been deposited in the Protein Data Bank under accession codes 6IBB and 6RNK. There are no restrictions on data availability.

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Sample size | No statistical methods were used to predetermine sample size. The determined sample size was adequate as the differences between experimental groups was reproducible, as indicated. X-ray diffraction data were collected until completeness of the data set. |
| Data exclusions | No data were excluded from the analysis. |
| Replication | All attempts at replication of biochemical and signaling assays succeeded. The experimental findings were reproduced in independent experiments. The number of independent experiments and biological replicates in each data panel is indicated in the figure legends.. |
| Randomization | No randomization was attempted or needed. Randomization was not formerly performed in this study as it did not involve animals and/or human research participants. |
| Blinding | Authors were not blinded. No blinding was attempted or needed. Blinding is not relevant for protein structure determination or functional assays as the results are not subjective. |

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ Antibodies |
| ☐ | ☒ Eukaryotic cell lines |
| ☒ | ☐ Palaeontology |
| ☒ | ☐ Animals and other organisms |
| ☒ | ☐ Human research participants |
| ☒ | ☐ Clinical data |

### Methods

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ ChIP-seq |
| ☒ | ☐ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |

## Eukaryotic cell lines

Policy information about cell lines

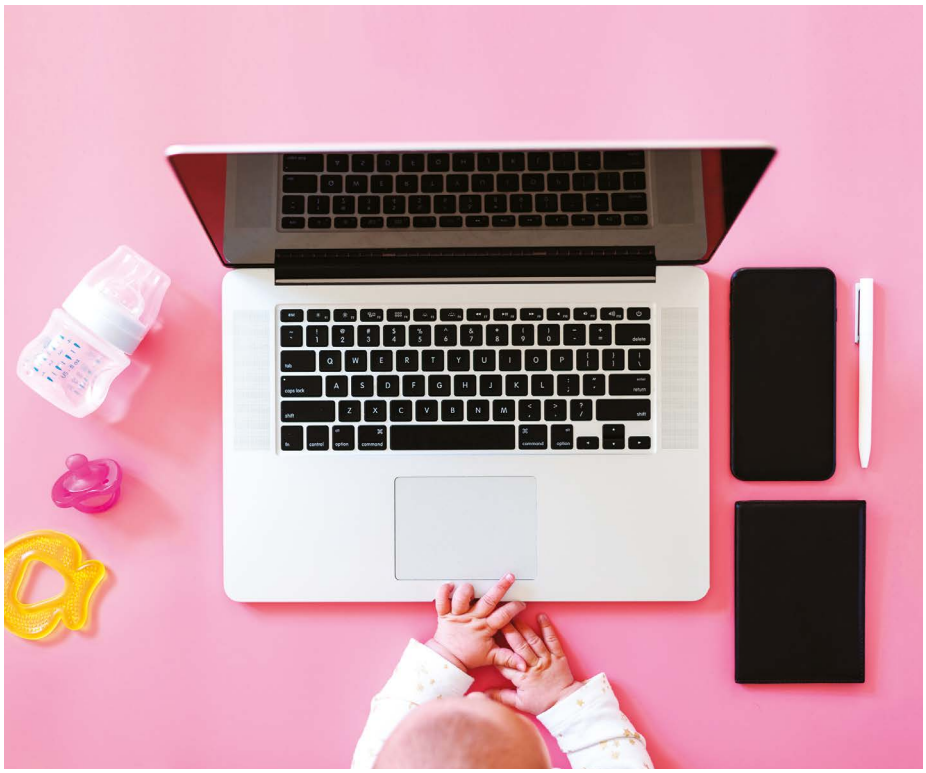
| | |
|---|---|
| Cell line source(s) | Sf9 insect cells (Life Technologies), HEK293FT (Life Technologies), SUCNR1 CHO-K1 (Novartis), S1P1R CHO-K1 (Novartis), SUCNR1 Chem1 (Millipore) |
| Authentication | Cell lines were maintained by the supplier. No additional authentication was performed by the authors of this study. |
| Mycoplasma contamination | Cell lines were tested for mycoplasma contamination and shown to be free from mycoplasma. |

| | |
|---|---|
| Commonly misidentified lines<br>(See ICLAC register) | No commonly misidentified cell lines were used. |

# Work

Careers online

**Why US parental-support policies need to improve: see go.nature.com/parent**



TROYAN/SHUTTERSTOCK

# SHARED PARENTAL LEAVE FOR THE FAMILY

## Lessons learnt in balancing academia and early parenthoood. **By Lynsey Bunnefeld**

My husband Nils and I work in the biological and environmental sciences department at the University of Stirling, UK, and we had our first baby in May 2018. Before our son Euan was born, we decided to make use of the United Kingdom's shared parental leave (SPL) policy. This scheme allows parents who meet certain eligibility criteria to share up to 50 weeks of leave, of which 37 are paid, in their child's first year of life. Our decision placed us among the 1% of all eligible couples nationwide who actually take the leave.

We had a loose plan: Nils would take two months' leave when our son was ten months old, at which point I would go back to full-time work. Because this would be in the final three months of our allotted leave, which in the United Kingdom are unpaid, and because my husband earns more than I do, it would involve a bigger salary loss than if I took those months off. However, we were able to take the financial hit, and although we reasoned that it might be difficult because Nils manages a large research group (mine, focused on ecology, is much smaller), we decided that it would be worth it for the time he would get to spend with our baby.

Fast-forward through the six life-changing months following Euan's birth, and it was becoming apparent that my mental health might benefit from my returning to work a little bit earlier than planned. Also, we were both concerned that Nils leaving his group to manage itself for two months might be asking too much – so we changed our plan. Under the policy, shared leave can be discontinuous, so we decided to split up the final three months of leave. One of us would work one week, while the other took leave – and the next week, we would switch. To minimize disruption to our departments, Nils committed to all of his teaching and administration during these three months. The leave was approved, and we were all set. Nils was excited about the time 'off' (I did try to tell him that a day with a baby is not really time off), and I was excited about activating parts of my brain that had been dormant for a while.

**"We were both exhausted and sleepwalking through our lives at work and home."**

This worked perfectly for some time, but in the third week of our 'one week on–one week off' cycle, trouble started to brew. Euan didn't nap, so Nils's Skype meeting with a collaborator couldn't happen. Students started knocking on my office door, asking why my husband hadn't replied to their e-mails. Nils read manuscripts in the evening once the baby was in bed. Assignment marking started to roll in. And things further unravelled from there. Although Nils continued to enjoy his time with Euan, he became increasingly anxious about work as he squeezed in e-mails and Skype calls whenever he could. He was not able to fully switch off his work brain and completely engage with the baby.

After a few more weeks, we adjusted our schedules so that we were both working part-time each week. This did alleviate some stress, and Nils stayed more on top of his responsibilities – but it was hard for me, having just returned to work, to get into any kind of rhythm. We were both exhausted and sleepwalking through our lives at work and home.

Many countries don't have policies similar to the United Kingdom's SPL, but here is our advice to academic couples who are in a position to make use of such benefits in the United Kingdom and elsewhere.

### Top tips

Make use of SPL. We affectionately call our SPL a 'car crash', but for Nils, Euan and I, the crash was totally worth it. Nils has a much better appreciation of what a day with a baby is like, the two had lots of fun together and the baby is totally happy at home with either me or my husband.

Actively put measures in place to ensure that the partner at home with the baby can be fully engaged with being at home. In hindsight, we agree that Nils should have made sure that all of his research students had alternative supervision during his leave, and he should have more clearly communicated to his research group

and network of collaborators that he was taking time off. He should also have declined at least 50% of requests for peer review and for help administering PhD-thesis defences. When the workplace does not support the parent on SPL fully, the other parent, usually mum, is also left unsupported and is unable to return to work as effectively.

Fight for appropriate coverage at work while taking SPL. Nils was on (unpaid) leave for 6 weeks over 3 months, but we estimate that he was working for about 80% of his 'normal' full-time hours rather than the 50% he was paid for during that period. Of course, this is partly due to his own conscientiousness — many scientists work far beyond their salaried hours, especially if they have a group that they feel responsible for. You might think that cover is not needed for such a short period — however, it absolutely is. That 6 weeks (or 30 days) of work needed to be done by someone. Our institution's policy is that time should be split 40:40:20 between research, teaching and admin. Assuming cover is required only for teaching and admin, we needed assistance for 18 days, either from someone inside the institution or from an external short-term contract worker. We suggest approaching human resources, the head of your department or your institution's equality and diversity committee before your leave begins in order to request this assistance.

**"Don't be too hard on yourself in the first few months with the baby."**

We can't be sure, but we think that one period of continuous leave might have helped matters. Our colleagues and Nils's research group might have found support elsewhere when they needed it, and it might have been easier for him to really switch off and be a stay-at-home dad for a short while.

Don't be too hard on yourself in the first few months with the baby. And once you're back at work, it takes a while to catch up with research — so enjoy having that time to think about non-baby subjects, get up to speed with new research and spend time with your colleagues. If you've implemented the tips above, you can be relaxed in the knowledge that your baby is at home having a ball.

Euan, now almost one and a half years old, is in a nursery three days per week, Nils is back to full-time work and I am adjusting to working part-time. Really, the challenge of navigating this new normal is just beginning. We'll not have a chance to relive our baby's first year, so we're hopeful that this post and our advice will help other new parents to get the most out of SPL.

**Lynsey Bunnefeld** is a lecturer in ecology and evolution at the University of Stirling, UK.

# YOU ARE NOT AN IMPOSTOR

## Ways to control the voice in your head that insists you're not good enough. **By Desiree Dickerson**

As I sit down to write this piece, a voice in my head tells me: "You can't do this," and "Who do you think you are?" Tension grows. Writing about well-being starts to stress me out. "This needs to be perfect," the voice continues.

This voice is not unique to me; we all have one. It is a product of our beliefs and our mindset. It influences how we perceive the world, our position in it and how we think, feel, act and interact.

It has driven many of us to academic accolades and career advancement — both measures of success according to most social standards.

But for some of us, this voice can denounce us as 'impostors' in academia and demand that we work twice as hard. Gradually, every day begins to feel like the morning of an examination. New ideas are dismissed with negative thoughts such as: "If I thought it, then it must be obvious." We read and reread to see how others have said what we want to say, because surely they said it better and more clearly. We silence our curiosity and don't speak up in lectures or meetings, missing invaluable learning opportunities.

The pursuit of excellence might have driven us to get high marks at university, but this perfectionism has become so ingrained that it fuels our need to forfeit rest as we work through the weekend. It underlies our tendency to amplify the criticism over the praise. We drag out deadlines as we search for something 'better' or 'more perfect'. Academia might benefit from this imbalance, but often our health as scientists does not.

Looking back, I can see that this voice played a large part in my departure from academia. Now that I run well-being and resilience workshops for academic institutions across Europe, and work one-to-one with academics as an academic-resilience coach, I know I am not alone.

After leaving academia, I decided to apply my skills as a clinical psychologist to change the narrative. First, I needed to dial down the fear and self-doubt that were so easily evoked in me.

To do that, I had to recognize the voice for what it was — a negative influence that I was allowing to make big life choices for me. I had to challenge the internal dialogue telling me I wasn't good enough, and to equip my new voice with arguments that recognized my strengths rather than magnified my fears. I realized that I had to develop a voice that could be compassionate in the face of setbacks — that would talk to me as I would talk to a good friend.

CAROL YEPES/GETTY

And, crucially, I needed to challenge the behaviours — avoidance, procrastination — that were empowering that voice and maintaining the cycle of self-doubt. These behaviours, of course, made me think that my old voice was right, that "I clearly wasn't good enough."

Writing a new script for the voice in my head is an ongoing process. I can't say I've killed off the character entirely, but it no longer plays the lead part. To complement the cognitive behavioural techniques that I used to rewrite my voice, certain specific, learnable exercises have helped me to gain more control.

I started practising mindfulness meditation to gain more control over where and how often my mind wanders. It helps me to be less emotionally reactive to things like criticism and feedback, less preoccupied by the progress of others and better able to focus on what I want to bring to the table. If you're interested, Mark Williams and Danny Penman's *Mindfulness: An Eight-Week Plan for Finding Peace in a Frantic World* (2011) was a good starting point for me.

I restructured my day to prioritize activities that make me most productive. I rate my sleep above all things and I exercise, no matter the deadlines, because I know it helps me to manage stress better, think more clearly and focus for longer (and it just makes me a much nicer human — to myself and to others).

By muting parts of that inner voice — the ones centred on perfection, worry, fear and guilt — you too can create space. Mental space and energy can be freed up to think, create, be present, ask questions, learn and relax. Imagine your life without that weight, without that constant pre-exam tension. Imagine academia without it.

**Desiree Dickerson** is a neuroscientist and a clinical psychologist (www.desireedickerson.com).

# Where I Work
# Gladys Ngetich

Photographed for *Nature* by Leonora Saunders.

I spend a lot of time with jet engines as part of my PhD, trying to improve them. This one is a Concorde engine — it's a demonstration model that's used to show where the various components of the engine are located, and it sits on the ground floor of the institute where I work.

Jet-engine turbines can reach temperatures of 2,000 °C. I'm exploring ways to use less air to cool them and to reduce carbon dioxide emissions. I test the performance of films that coat different engine-blade designs in a one-metre-long research tunnel at one of the 20 research facilities at the institute.

As a PhD student, I get limited time with the rig, which I use to test whether our models are accurate. A mock turbine is wired with small, flexible tubes to monitor pressure as well as with tiny temperature sensors.

The set-up leaves just enough space to sit at a computer table to record data.

The machine is deafening and I often do experiments that involve ultraviolet light. Everyone has to wear ear protection and goggles. To ensure that my experiments run smoothly, I prefer to process and analyse data early in the morning, late in the evening or at weekends, when colleagues are largely absent. I've found that the small breakthroughs always come over the weekend.

Since I left Kenya to start my PhD four years ago, I've relied on a strong support system during tough times. Some personal rituals help me to focus. I play high-energy dance music — including Bongo, or Swahili hip hop.

Getting my brain into a rhythm, just like when I'm running, helps me to relax and solve problems when something is not working.

I also like to keep my work area simple and clutter-free to avoid distractions. I don't have any posters or any other personal touches. It's just me, my tunes and the rig.

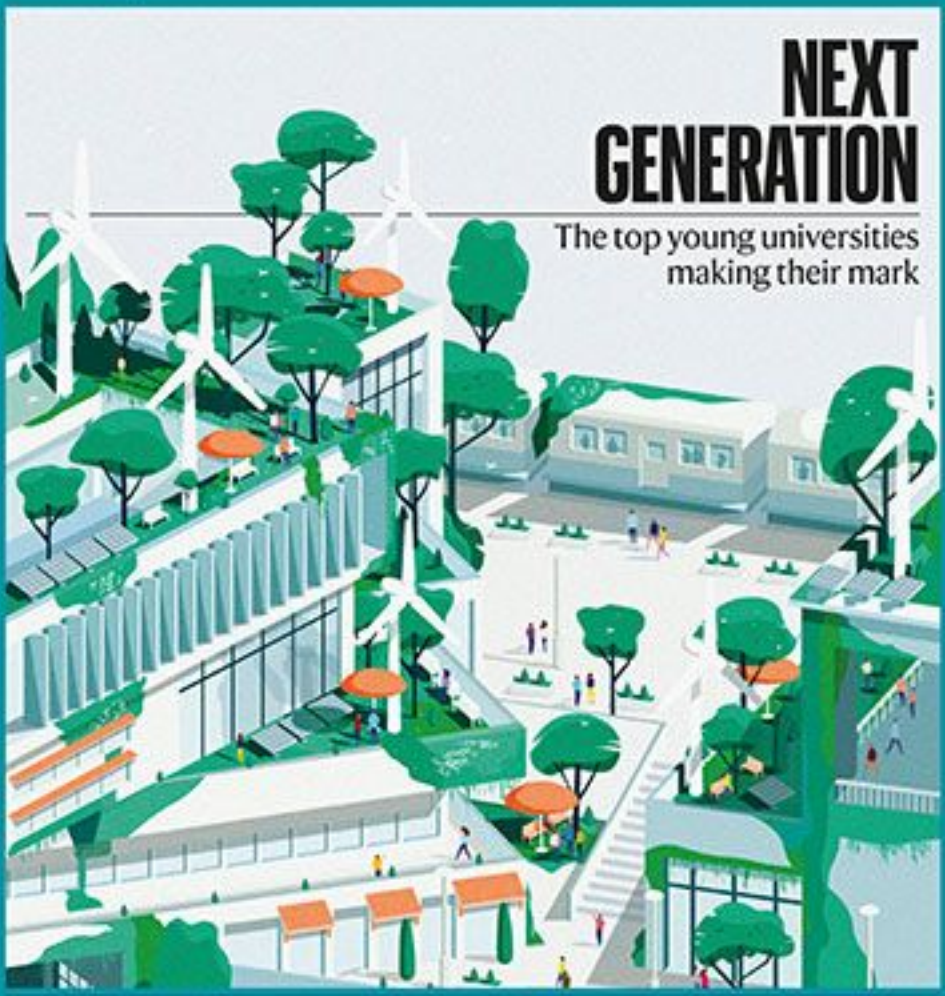
Oddly, as my experiments wind down, I feel like I've developed Stockholm syndrome. I've spent so much time here, yet I realize that I've formed some kind of attachment to this place — and I know that when I ultimately leave, I'll miss it.

**Gladys Ngetich** is a PhD student at the Oxford Thermofluids Institute at the University of Oxford, UK. **Interview by Virginia Gewin.**

nature

# index

## Young universities




**Fountains of youth**
Output graph shows strongest countries

**Learning curves**
The higher education commodity boom

**Rapid risers**
Brief histories and big futures

nature
index

# Young universities



**Nature Index Young universities 2019**
Nature Index Young universities 2019, a supplement to *Nature*, is produced by Nature Research, the flagship science portfolio of Springer Nature. This publication is based on data from the Nature Index, a Nature Research database, with a website maintained and made freely available at natureindex.com.

**Nature Editorial Offices**
The Campus, 4 Crinan Street,
London N1 9XW, UK
Tel: +44 (0)20 7833 4000
Fax: +44 (0)20 7843 4596/7

**Customer services**
To advertise with the Nature Index, please visit natureindex.com, or email clientservicesfeedback@nature.com


It takes time to build research institutions of quality and substance, but getting the right components together at the outset enhances the chances of success. The most successful among the Nature Index Young universities, defined as being aged 50 years or younger, are making remarkable headway in attracting talent and rising up the ranks through high-quality research outputs and collaborations.

The leaders of these higher-education institutions often cite similar reasons for their success. Many feel liberated from the traditions that characterize older institutions, and they list strong interdisciplinary cultures, a track record of innovation and the capacity to attract a more diverse student population.

The proliferation of new universities in the 1950s and 1960s has had remarkable effects on countries such as South Korea, where an explosion of higher-education opportunities has seen the proportion of 25–34-year-olds with a tertiary education surge from less than 2% at the time of independence in 1945 to 70% in 2017, among the highest rates worldwide. China and Singapore have also benefited from their commitment to revamping the research and education landscape.

The institutions featured in this supplement are outstanding performers in terms of the Nature Index metrics of article count (AC) and fractional count (FC). The first (AC) is the total number of articles published by an institution's affiliated authors in the 82 publications tracked by the Nature Index. FC measures the share of those institutions' contribution to each article.

The institutions pride themselves on promoting creative thinking, and offering leadership opportunities to young- and mid-career researchers who are encouraged to pursue unconventional research that sparks invention. As Christopher Barner-Kowollik, a macromolecular chemist at Queensland University of Technology in Australia, puts it: "Innovation occurs at the flanks of research, not the mainstream."

And yet, just as industry start-ups often struggle to make it past the first five years, young universities have an uphill battle ahead of them once the initial cash injections to establish them cease. Only those that have built reputations to rival their older peers will survive.

**Bec Crew**
Senior editor

**On the cover**
The future imagined: young universities embrace integrative technology, and a green ethos.

## Contents

# index


**Grégoire Courtine and Léonie Asboth, whose studies on rats at EPFL led to an implantable device to help paraplegic patients.**

# FAST LANE TO THE FUTURE

Unconstrained by centuries of convention, in the race for solutions these high performers are setting the pace in diverse research fields.

## SWISS FEDERAL INSTITUTE OF TECHNOLOGY LAUSANNE

2018 FC: 219.92 | AC: 542

Faculty: 4,700* | Students: 11,134

PhD graduates: 2,216

An implantable device that has restored the ability of three patients with paraplegia to walk is one of the most promising innovations in development at the Swiss Federal Institute of Technology Lausanne, Switzerland (EPFL).

The wireless implant, made up of an array of electrodes stretched over the spinal cord, targets individual muscle groups in the legs to mimic the signals fired in the brain when walking.

Volunteers in the clinical trial, David Mzee, Gert-Jan Oskam and Sebastian Tobler, have endured months of training and physical therapy to regain voluntary control over their leg muscles after several years of paralysis. They are now able to walk with the aid of crutches or a walker.

The study, published in two papers last year in *Nature* and *Nature Neuroscience*, is led by Grégoire Courtine, a neuroscientist at EPFL's Brain Mind Institute, and Jocelyne Bloch, a neurosurgeon at the Lausanne University Hospital (CHUV). The researchers saw continuous improvements in the patients' motor function, even after the device was switched off.

The research follows a study published in early 2018, led by Courtine's colleague, Léonie Asboth, which produced similar results in paralysed rats. The team observed for the first time how the brain can reroute motor commands through alternative pathways to the spinal cord.

"During a thesis, we all wonder at some point if what we're doing is going to have an impact," says Asboth. "Being able to see the first implications of this research on patients with spinal cord injury was very rewarding for all of us."

An EPFL spin-off company, GTX medical, co-founded by Courtine and Bloch, is now developing the technology for use in hospitals and clinics. EPFL is exclusively licensing the patents to GTX medical, while also hosting many of its 40 researchers and clinicians.

Known formerly as the École polytechnique de l'Université de Lausanne, EPFL was established in 1969 as a university in its own right following a decision by the Swiss parliament to create a second federal institute of technology in addition to ETH Zurich. It is the third-highest ranked young university in the Nature Index and Switzerland's only representative among the leading 100 young universities. **Bec Crew**

*includes technical staff

## SHANGHAITECH UNIVERSITY

2018 FC: 36.34 | AC: 164

Faculty: 521 | Students: 3,165

PhD graduates: 41

China's ShanghaiTech University's high-quality research output has grown rapidly since its foundation in 2013. It is the world's fourth-fastest rising young university and is ranked 22nd among the Nature Index Young universities.

Ning Zhijun, assistant professor in the School of Physical Science and Technology, says he was attracted to ShanghaiTech's systematic materials science research, through which he and his team have produced high-profile papers in *Nature*, *Nano Letters* and *ACS Nano*.

EPFL/ALBAN KAKULYA

UNIST

**Hye Young Cho (left) and Young Shin Yoo at UNIST focus on new cells that are safer than conventional lithium-ion batteries.**

"Unlike established institutions, ShanghaiTech as a new university does not impose heavy publication pressure on young scientists," he says. "Our university offers generous research funding, so we can take on original studies that might not be published in the short term."

Zhong Chao, a nanomaterials scientist in the physical science school, says that ShanghaiTech differs from older Chinese universities because its young scientists are not required to join teams of senior scientists when they are first recruited. This means they are free to explore their own research interests with fresh eyes.

"In a new university like ShanghaiTech, there aren't so many established figures, so young scientists can independently take on some risky research, which is potentially more innovative," says Zhong. Peer collaboration is also easy, he says, unencumbered by the need to consider seniority.

ShanghaiTech was jointly launched by the Shanghai municipal government and Chinese Academy of Sciences (CAS). As is the case with other young universities, it was not immediately permitted by China's Ministry of Education (MOE) to recruit doctoral students in its own name, but CAS as a co-founder has helped it overcome this difficulty.

"CAS recruited doctoral students for us in the name of its Shanghai branch, and then transferred these students to us," says Zhong. In 2018, the MOE allowed the university to independently enrol its first doctoral students in materials science and engineering. **Hepeng Jia**

## ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

2018 FC: 68.88 | AC: 161

Faculty: 322 | Students: 5,272

PhD graduates: 116

Sang-Il Seok has a vision for his lab's perovskite solar cells: covering the decks of a crude-oil tanker and supplying clean power to the vessels that haul the dirtiest of fuels. It's a juxtaposition familiar to his university, the Ulsan National Institute of Science and Technology (UNIST), which is the tenth-highest ranked young university in the Nature Index.

The city of Ulsan is known as an industrial hub of South Korea, but since its establishment in 2009, UNIST has earned a reputation for clean energy research, including batteries and solar cells.

"UNIST has become one of the strongest places to study green technology in Korea," says Seok, a materials scientist whose research group specializes in solar cells that use a perovskite compound as the light-absorbing layer, which is easier to fabricate than more common materials such as silicon.

In 2017, Seok and his colleagues set the benchmark in perovskite-solar-cell efficiency of 22.1%. That benchmark has since been raised to 25.2%, just behind silicon at 27.6%.

Seok's early success at UNIST was bolstered by funding for his lab of more than one billion South Korean won (US$826,000) from the university, part of an initial investment in its new campus.

UNIST's focus on building a workforce skilled in new energy technologies is more important to the local community than ever, as the Korean shipbuilding industry, which was booming when the institution was first established, has suffered a significant downturn in recent years. "We need to educate the students who can get these good jobs in the near future," says Seok.

Established to challenge South Korea's top technical universities, UNIST mandated that all courses are to be taught in English to boost its international competitiveness. This is crucial for researchers in its Fluidics and Reactions Using Integrative Technology and Science (FRUITS) Lab, for example, who collaborate with EPFL in Switzerland, the National University of Singapore, and several teams in the United States on cell-to-cell communication, lab-on-a-chip technology, and nanodevices for use in medical research.

"There is no borderline in science," says FRUITS Lab group leader, Yoon-Kyoung Cho. "Many of my students have learned self-confidence through active international collaborations. We have all our meetings in English, which was not easy in the beginning, but it becomes so natural that the students ask good questions at big international conferences, which makes me proud." **Mark Zastrow**

## SHENZHEN UNIVERSITY

2018 FC: 52.48 | AC: 179

Faculty: 3,647 | Students: 34,156

PhD graduates: 26

Following Shenzhen University's (SZU) establishment in 1983, two of China's top universities, Peking and Tsinghua, seconded teaching staff to the fledgling institution. The move was to support the city of Shenzhen's development as one of four 'special economic zones' in southeastern coastal China, which were created in 1980 to attract foreign investment and technology.

High-profile alumni who have since cut their teeth at SZU include computer scientist, Ma Huateng, who is the founder and chief executive of the Chinese social media behemoth, Tencent, and software engineer, Shi Yuzhu, who set up the online gaming company, Giant Interactive Group. Alumna Tu Hongyan, now chairperson of Hangzhou-based silk brand, Wensli, was named one of Forbes' top Chinese women in business in 2018.

Today, one of SZU's most highly cited scientists is Zhang Han, a professor of optics and photonics. Han uses graphene and other two-dimensional materials to create laser photonics devices, which have applications in fields such as medicine, communications and quantum information science.

"Our research received heavy investment from the Shenzhen municipal government, which considers new materials as one of its priority high-tech industries," Han told Nature Index.

Han joined SZU in 2013 as a 'young thousand-talent', part of China's Thousand-Talent scheme, launched in 2008 to attract leading scholars. He says SZU's advantages over more established universities in China include the encouragement it offers to young scientists who are keen to pursue new research areas, and its strong support for international collaboration.

**"It was rather mundane biology. I wanted a more meaningful experiment."**

The university has partnerships with 256 universities overseas for collaborative research and student training. It also has strong links to local industry due to Shenzhen's status as a high-tech hub.

Last year the university's total research budget exceeded 1.1 billion yuan (US$153 million), up from 100 million yuan in 2013, and it received 302 grants from the National Natural Science Foundation of China.

SZU is the third-fastest rising young university in the world and is ranked 13th among the Nature index Young Universities. It is also the third-fastest rising young university in the fields of chemistry and physical sciences. **Hepeng Jia**

## DAEGU GYEONGBUK INSTITUTE OF SCIENCE AND TECHNOLOGY

2018 FC: 19.99 | AC: 53

Faculty: 260 | Students: 1,449

PhD graduates: 26

When robotics engineer, Hongsoo Choi, visits his neuroscience collaborators to work on their medical microrobots, he doesn't even have to walk outside. All six departments of the Daegu Gyeongbuk Institute of Science and Technology (DGIST) are housed within a tight cluster of buildings constructed in 2010. "It helps a lot, actually," he says. "I can just stop by for discussions if I need. That's a big advantage."

The layout reflects one of the South Korean university's core principles: 'convergence', its preferred term for an interdisciplinary mindset in research and study.

In May 2019, the team of roboticists, engineers and neuroscientists reported that they had created sphere- and helix-shaped microrobots that can deliver transplanted stem cells inside a live mouse. When a magnetic field is applied, the bots can roll along the walls of blood vessels or swim through fluids, carrying stem cells to their target.

In proof-of-concept experiments six years earlier, Choi guided microbots as they swam around plastic containers carrying kidney cell cultures. It was remarkable robotics, but rather mundane biology. "I wanted a more meaningful experiment," he says.

This led him to the neuroscientists next door, who could culture neural stem cells from the hippocampi of mice, part of the brain involved with memory, learning and emotion. The study, published in *Science Robotics*, reports that the microbots could carry these tiny payloads until they differentiated into several types of brain cells. They also navigated the arteries of a dead rat's brain, demonstrating the potential to deliver therapeutic cells to targeted areas to potentially restore brain functionality.

Choi's team involves a number of young researchers with diverse expertise in robotics, such as Junhee Choi, who in a separate project helped to develop an ultrasonic device for root-canal treatment, and PhD candidate, Eunhee Kim, who focuses on the adhesion properties of microrobots in regenerative medicine.

DGIST, which was established just 15 years ago, actively encourages cross-disciplinary projects, as they have a better chance at being awarded internal grants, says Choi. "It's not a rule, but it's the culture." The university is the seventh-fastest rising young university and is ranked 50th among the top young universities in the Nature Index. **Mark Zastrow**

## OREGON HEALTH AND SCIENCE UNIVERSITY

2018 FC: 55.48 | AC: 151

Faculty: 2,900 | Students: 4,706

PhD graduates: 28

Brian Druker, director of the Oregon Health and Science University's (OHSU) Knight Cancer Institute in Portland, in the United States, is developing targeted treatments for acute myeloid leukaemia (AML), the most common form of blood cancer in the United States. This highly lethal disease affects the myeloid cells in bone marrow and kills more than 10,000 people in the United States annually.

Over the past 40 years, progress towards new, more effective treatments for AML has been slow. "There have been few changes in the way this form of leukaemia has been treated," says Druker.

# index

QUT

**QUT nanotechnologist, Jennifer MacLeod, is studying how to grow and modify two-dimensional materials.**

Two decades ago, Druker was involved in the development of the first targeted drug for chronic myelogenous leukaemia, a slow-growing form of leukaemia. The treatment, marketed as Gleevec, transformed the disease from a life-threatening illness to a manageable condition, with 90% of patients living for at least five years after diagnosis.

In 2016, Druker and his colleagues established the Beat AML programme, a long-term collaborative clinical trial that aims to uncover targeted treatments for various forms of acute myeloid leukaemia.

As part of this work, the team generated the largest data set of its kind, drawn from 672 tumour samples from 562 patients. The findings, published in *Nature* in October 2018, will help researchers pinpoint which genetic markers are sensitive or resistant to treatment, to inform future clinical trials.

OHSU, the former University of Oregon Health Center, became independent of the university in 1974, and was renamed in 1981. It is now the largest employer in Portland, and the only academic health-care centre in the state, comprising three campuses, two hospitals and numerous clinics.

OHSU is the most prolific young university in life sciences research output in the Nature Index, and is ranked 11th among young universities in the Nature Index.

Working at a relatively young institution has enabled Druker to "move quickly and get things done without having to navigate layers of bureaucracy". Having the freedom to think outside the box is another advantage. "Being willing to embrace ideas that might be considered outside the mainstream has allowed us to develop new and paradigm-changing research," says Druker. **Gemma Conroy**

## QUEENSLAND UNIVERSITY OF TECHNOLOGY

2018 FC: 27.85 | AC: 117

Faculty: 2,110 | Students: 47,592

PhD graduates: 327

A synthetic material that can shift its structure under different light conditions has been developed by macromolecular chemist, Christopher Barner-Kowollik, at the Queensland University of Technology (QUT) in Australia.

When the material is exposed to green LED light, its chemical bonds strengthen to produce a hard, stable structure. In darkness, the material transforms into a soft, liquefied mass. Not only is the material reprogrammable, but it's inexpensive to produce, consisting of just two chemical compounds. One of them, naphthalene, is an active ingredient in moth repellents.

This light-stabilized dynamic material could be used as a 3D-printing ink for creating temporary scaffolds that support free-hanging structures, which are notoriously difficult to print using current methods.

In an effort to advance understanding of the material, Barner-Kowollik and his team published in the *Journal of the American Chemical Society*, in June 2019, rather than patent it, in the hope that other teams will explore its potential. They are part of an international collaboration involving Ghent University in Belgium and Karlsruhe Institute of Technology, Germany.

The QUT researchers are based in the university's science and engineering faculty, which is one of the largest university faculties in Australia. They work alongside nanotechnologist, Jennifer Macleod, who uses scanning probe microscopy and X-ray photoelectron spectroscopy to investigate how to grow and modify two-dimensional materials, such as graphene.

QUT was established just 30 years ago, after operating for 20 years as the Queensland

Institute of Technology. It's the sixth-fastest rising young university, as tracked by the Nature Index, and is ranked 30th of the young universities in the index.

Whether it's designing a new laboratory space or finding new collaborators, Barner-Kowollik says he's able to effect change quickly at QUT, and has been given the freedom to explore new kinds of research questions. "Innovation occurs at the flanks of research, not within the mainstream," he says.

QUT has two campuses in Brisbane, and in 2013, it opened the Science and Engineering Centre at its Gardens Point campus. In addition to teaching spaces and educational facilities, the centre houses the Institute for Future Environments, which brings together more than 300 scholars from different fields to collaborate on large-scale projects relating to natural, built and digital environments. **Gemma Conroy**

## UNIVERSITY OF PARIS-SUD

2018 FC: 71.08 | AC: 574

Faculty: 4,300 | Students: 31,800

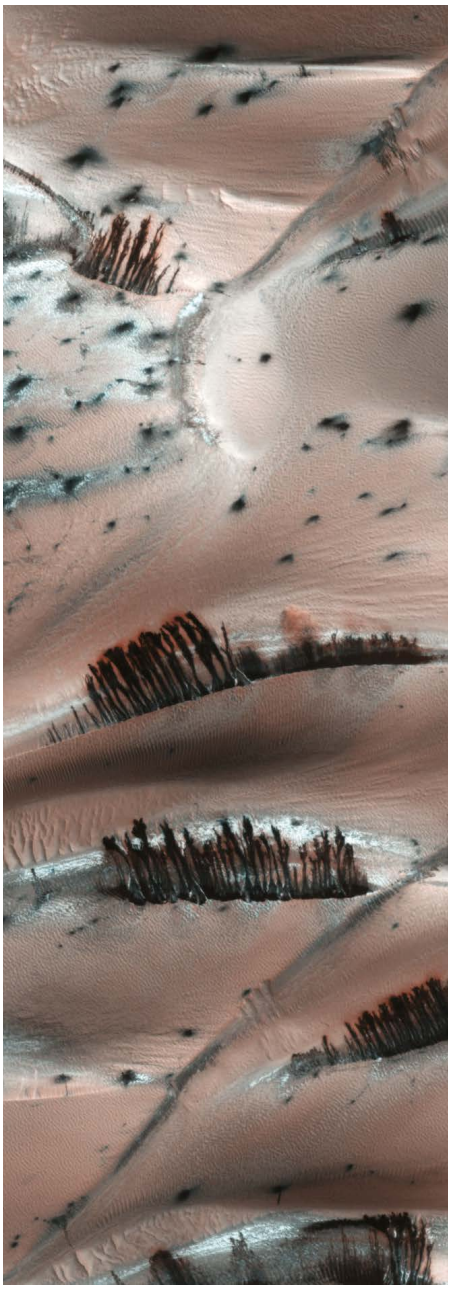
PhD graduates: 688

François Costard, a geomorphologist at the University of Paris-Sud, is at the forefront of work investigating surface features on Mars for historic evidence of oceans. His latest paper, published earlier this year in the *Journal of Geophysical Research: Planets*, suggests that the Lomonosov crater in the planet's north could have been the source of a mega-tsunami three billion years ago.

The scenario involves an asteroid collision of similar impact to the one that wiped out the non-avian dinosaurs on Earth 66 million years ago. In the case of Mars, it's thought that the asteroid slammed into a shallow ocean, causing a massive wave to form. The research provides evidence that liquid water could have persisted on Mars for millions of years.

"This has implications for the total inventory of water on Mars, how it evolved, and the potential for the origin and survival of life on the red planet," says Costard, director of the planetary geomorphology team at the University of Paris-Sud and director of research at France's National Center for Scientific Research (CNRS).

He notes the advantages of research at a young institution like the University of Paris-Sud: "The youth of our institution favours the possibility of us having young research scientists and financial support for new, especially interdisciplinary, programmes."



HIRISE/MRO, NASA/JPL/UNIVERSITY OF ARIZONA

**A true alien landscape, these surreal sand dunes were photographed by the Mars Reconnaissance Orbiter near the red planet's north pole.**

Originally part of the University of Paris, Paris-Sud was established as a university in its own right in 1970, and now has several campuses in the southern suburbs of Paris, including its main campus in Orsay. It is ranked ninth in the young universities in the Nature Index, and its highest subject rank is in the physical sciences, where it is also placed ninth.

In 2014, the University of Paris-Sud was a founding member of the University of Paris-Saclay, a 'mega-university' that brings together 19 universities, colleges and research centres in the south of Île-de-France. By 2020, when the University of Paris-Sud will be officially integrated into the University of Paris-Saclay, the consolidated institution will represent 15% of France's total research output. **Bec Crew**

## HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

2018 FC: 108.39 | AC: 310

Faculty: 680 | Students: 11,205

PhD graduates: 273

Despite being a relatively small institution, the 28-year-old Hong Kong University of Science and Technology (HKUST) has been consistently ranked among the world's top young universities, noted for its growth and strong reputation.

In the Nature Index, HKUST is the fifth-highest ranked young university. It was ranked 32nd in the QS World University Rankings 2020, published by UK education company, Quacquarelli Symonds, and came second in its 50 Universities Under 50 ranking. It took the top slot in the 2019 *Times Higher Education* Young Universities Ranking for the second year running.

In addition to its traditional areas of strength – computer science, quantum physics and medical sciences – HKUST has achieved significant progress in molecular neuroscience, where researchers are investigating the proteins that promote the development of neurons and the mechanisms underlying neurodegenerative diseases, such as Alzheimer's.

HKUST's vice-president of research and development, Nancy Ip, says that the university's flexible and efficient decision-making and its ability to define its own traditions and swiftly adapt to challenges have brought international recognition. She says the university's location is also an advantage, facilitating collaboration with international and Chinese universities.

The career path of Qian Zhang, a computer scientist who joined HKUST in 2005 from Microsoft Research Asia, is an example of such collaboration. She has connected with a number of industrial partners, including Microsoft and Intel, to develop new wireless connection technologies. In 2009, she established a joint lab with Chinese telecoms giant, Huawei.

Zhang is the inventor of more than 50 granted and 20 pending international patents, and in 2016 became the university's youngest endowed chair professor. "Compared with other more established universities in Hong Kong, the key to HKUST's success is its openness," she says. **Hepeng Jia**

# Perspective: Passing the baton

## New universities with broad focus are spreading knowledge globally. **By Philip G. Altbach**

In the past half-century, the higher-education sector has mirrored the patterns of many luxury retailers: it has embraced massification and extended its product range to a wider market through the proliferation of new participants. The number of new academic institutions established worldwide in recent decades is unknown but undoubtedly runs into many hundreds.

The vast majority of these new universities would not appear in the Nature Index, which measures high-quality research outputs, or in the Academic Rankings of World Universities (Shanghai rankings) or *Times Higher Education*'s World University Rankings. This is because most young universities are local institutions focused on teaching young people, many of whom are the first in their families to attend a post-secondary institution.

Indeed, the rankings, including Nature Index's, overemphasize the elite research-intensive sector of post-secondary education. Teaching excellence is undervalued or ignored, in part because measuring quality is difficult. Research productivity has traditionally been seen as most prestigious.

Since 2000, the number of students in higher education has more than doubled, exceeding 210 million globally, with the majority of this expansion taking place in developing and middle-income countries. Where governments have been slow or unable to invest in expansion, the private sector has taken over, bringing tremendous variation to young universities. Private universities now represent the fastest growing segment of post-secondary education worldwide.

Most of the young universities profiled by Nature Index for their impressive research performance are public institutions in Asia, Australia and Europe. It is significant that none are in Africa or Latin America, where higher-education investment and quality has lagged behind the rest of the world. In North America, the new universities that have achieved excellence are now not so young, having mostly been established in the expansion period of the 1960s. Institutions such as the University of California, San Diego, the State University of New York at Stony Brook and the University of Waterloo in Canada are good examples.

The Nature Index's interest, of course, lies with STEM-focused institutions, and it is worth noting that recent major investments in higher education have been made in universities that are strong in science and technology. Publishing in STEM fields typically yields more citations, so raises the global visibility of both the researchers and their universities. But some excellent young universities that focus more broadly have been established in the past two decades.

JACOBS UNIVERSITY

**The number of students in higher education has more than doubled since 2000.**

> **Teaching excellence is undervalued or ignored, in part because measuring quality is difficult.**

**Philip G. Altbach,** research professor and founding director at the Center for International Higher Education, Boston College. e-mail: philip.altbach@bc.edu

In India, key examples are the Shiv Nadar University and O. P. Jindal Global University, both near Delhi, and Azim Premji University in Bengaluru, which focuses on research and training in education. These private non-profit universities, founded by philanthropists with deep pockets, are pioneers in their governance, curriculum and orientation compared with other Indian universities.

In the United States, the Olin College of Engineering, established in 1997 with a US$460-million grant from the Olin Foundation, aims to revolutionize undergraduate engineering education by crossing disciplinary boundaries and eliminating lectures. Jacobs University in Bremen, Germany, is another 'new model' university, functioning entirely in English, with a majority of international students.

These dynamic institutions illustrate a growing trend around the world, perhaps harking back to the early twentieth century in America, where the Rockefellers (University of Chicago) and Stanfords (Stanford University) were bankrolling new universities with new ideas.

The most successful young universities, whether public or private, are characterized by significant investments, innovative ideas about governance, curriculum and social responsibilities, and forward-thinking leadership. Whether all will succeed in the long run, as practices become entrenched and funding may dwindle, is unclear, but these impressive young institutions are beacons for the future of higher education.

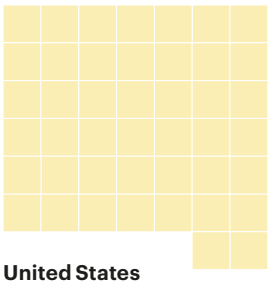
# Putting youth on the map

Global research ranks are reordered when visualized according to the output of young universities in the Nature Index.

## HOW TO READ THIS GRAPH

Each square represents a fractional count (FC) rounded to five. Fractional count measures the share of authorship of each article. Data are for 2018. Colour represents the young universities' contribution to each country/region's total research output in the Nature Index.


## THE AMERICAS



**United States**
221.29 (1%)

The United States has the highest output in the Nature Index overall, but the contribution of its young universities is relatively low. The Oregon Health and Science University (11th), the University of Texas at Dallas (17th) and the University of Alabama at Birmingham (25th) are its best-performing young universities for high-quality research output in the natural sciences.

**Brazil**
29.71 (10%)



**Canada**
17.33 (1%)

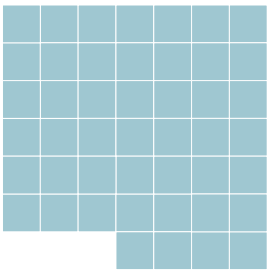
**Chile**
12.06 (12%)

## EUROPE



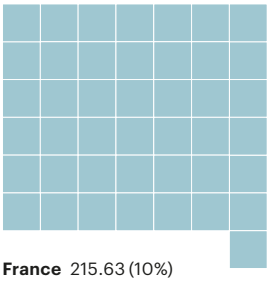
**Germany** 240.20 (5%)

Germany has 11 of the top 100 young universities in the Nature Index, the same number as China, but the collective output of Germany's young universities is less than one-third of their Chinese peers.

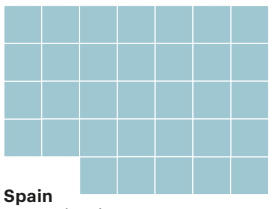


**Switzerland** 227.90 (16%)

EPFL is the only Swiss university among the top 100 young universities in the Nature Index.



**France** 215.63 (10%)



**Spain**
165.24 (15%)



**Italy** 94.20 (9%)



**Sweden**
53.84 (9%)

**Norway**
51.95 (27%)

**Austria**
51.92 (15%)



**United Kingdom** 37.37 (1%)

**Portugal**
32.42 (20%)

**Ireland**
17.20 (16%)



**Russia**
15.75 (3%)

**Greece**
12.21 (16%)

**Luxembourg**
12.02 (62%)

**Netherlands**
11.90 (1%)

**Czech Republic**
10.34 (5%)

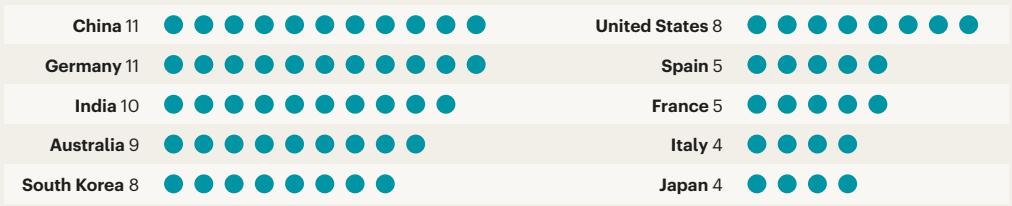
**Denmark**
8.55 (2%)

**Finland**
6.73 (3%)

**Cyprus**
4.98 (54%)

## WORLD LEADERS

China and Germany have the greatest number of young universities among the 100 young leaders in the Nature Index. Saudi Arabia and Singapore's few young institutions contribute the most to their countries' output.

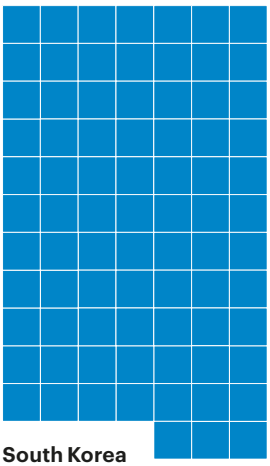



## ASIA PACIFIC AND MIDDLE EAST



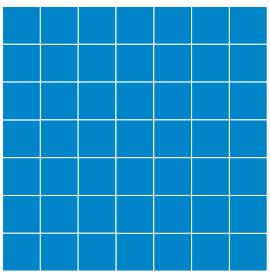
**Mainland China**
841.47 (8%)

The University of Chinese Academy of Sciences is the highest ranked young university in the Nature Index, with more than three times the article count of the nearest competitor (NTU Singapore). The top four fastest-rising young universities in the Nature Index are in China.



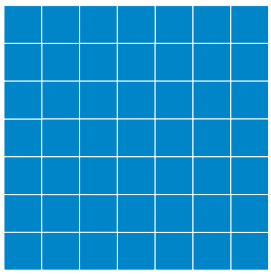
**South Korea**
401.04 (30%)

China has the highest fractional count (FC) of any country, as contributed by young universities. South Korea comes in second, with less than half of China's FC.



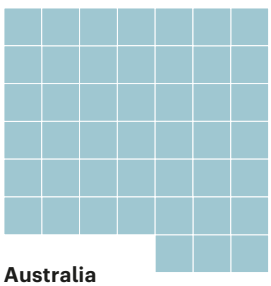
**Singapore**
245.69 (41%)

Nanyang Technological University with an FC of 232.51, is ranked number two among young universities in the Nature Index. Singapore is one of several countries with two institutions in the young universities top 100, including Norway, Austria, Israel, Sweden, Brazil and Portugal.



**India**
244.56 (26%)

The Homi Bhabha National Institute with an FC of 42.31 is the highest-ranked young university from India at 16th place overall, and the second-highest young graduate university. IISER Pune is second among Indian young universities at 20th place and IISER Bhopal was third in 26th position overall.

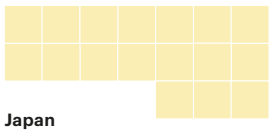


**Australia**
226.15 (18%)

Curtin University is Australia's highest ranked young university, at 23rd. Its closest competitor, the Queensland University of Technology, in 30th place, is the sixth-fastest rising young university in the world in change in FC between 2015 and 2018.

**Saudi Arabia** 104.77 (80%)

The King Abdullah University of Science and Technology at sixth is the highest ranked young graduate university in the index. It contributes a far greater proportion of total research output than its young counterparts in any other country.



**Japan**
85.46 (3%)

**Israel**
57.97 (10%)



**Taiwan**
32.41 (9%)



**Turkey**
12.30 (18%)



**Iran**
10.21 (9%)

**Thailand**
5.46 (14%)

**Qatar**
5.26 (44%)

SOURCE: NATURE INDEX

Q&A

# Subra Suresh
# Force for change

**Ranked second among Nature Index's Young universities, Nanyang Technological University is moving at an accelerated pace.**

Nanyang Technological University (NTU) is one of Singapore's top research institutes, and in recent years has emerged as a global leader in driving the 'fourth industrial revolution', a period defined by disruptive technologies such as the Internet of Things, robotics, virtual reality and artificial intelligence.

Established in 1991, and now the second-most prolific young university in the Nature Index, NTU has climbed the global rankings in research output and reputation.

Nature Index spoke with its president, Subra Suresh.

**How does NTU seek to engage with industry?**

NTU is engaging with some of the top industries from around the world. British jet-engine manufacturer, Rolls-Royce, for example, has partnerships with 29 universities globally and their largest partnership is with NTU. This year we renewed a five-year contract with them, worth $88 million Singapore dollars (US$63.5 million) to look at next-generation aircraft engines, 3D printing, digital manufacturing and many other topics.

Chinese retail and e-commerce company, Alibaba, established a joint AI research institute with NTU on our campus. It involves 25 Alibaba employees working with 25 professors here. And last year, American software company, Hewlett-Packard (HP), established its largest university partnership with NTU, with $84 million Singapore dollars in funding over four years for digital manufacturing technologies.

We also collaborate with Volvo. We converted its electric buses into fully self-driving vehicles, which are now being piloted on our campus. We're working with Singapore's Land Transport Authority and other government organizations to explore different types of autonomous vehicles.

We have more than 180 companies from around the world on campus. These provide opportunities for faculty and students to connect research and education to industrial practice, and job opportunities for our students once they graduate.

**How has NTU's location contributed to its growth?**

China and India, the two most populous countries in the world, are culturally represented in the Singaporean population, which attracts Indian and Chinese students because of cultural affiliations, geographical proximity and the familiarity of living in Asia. We have more than 20,000 NTU alumni who occupy prominent positions in China today. This representation gives us a natural connection to China.

Indonesia and Malaysia are also close to us and have historical ties to Singapore. Many top US companies have regional headquarters in Singapore, including HP and Procter & Gamble. This helps us to connect with industry.

English is the primary language of Singapore, and our primary language of instruction. We have all the practices, policies and procedures of a Western university. Being located at the crossroads of Asia as a multicultural, multiracial society, but with a very strong Western focus, makes us unique.

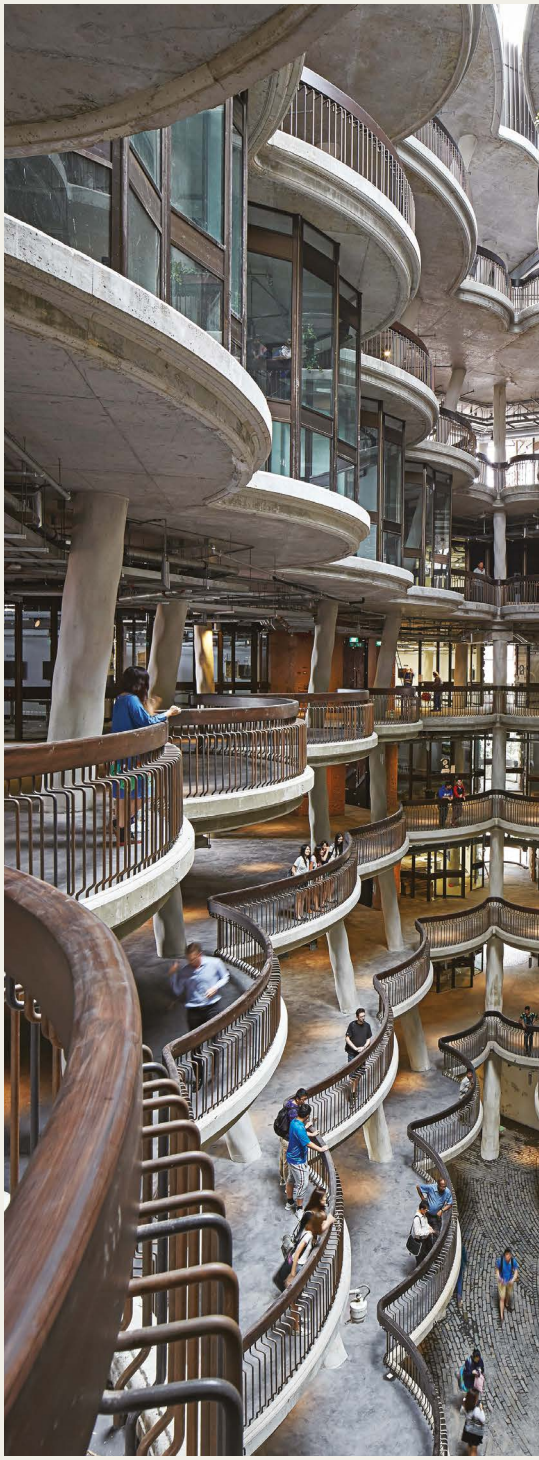
**NTU was founded in 1991. How has it grown in such a short amount of time?**

Twenty years ago, the Singapore government took a long-term view on the importance of having world-class universities here, making Singapore a destination for academic, industrial, entrepreneurial and innovation talent.

They created new funding to make this possible and invested in university buildings, labs and facilities. At the same time, the Economic Development Board of Singapore aimed to attract companies to do high-end research.

**What advantages do young universities have?**

Six years ago, NTU set up the Lee Kong Chian School of Medicine in partnership with Imperial College London. We developed



**More than 180 companies on NTU's campus connect research to industrial practice.**

a modern curriculum using the latest technology, online learning and continuous assessment. We're using virtual and augmented reality to teach subjects such as cardiology and anatomy. That's hard to do in an established medical school because you have to retrain your medical doctors and professors before you can educate your students.

It's also easier for a new medical school to leapfrog old technologies and equipment and go straight to the latest ones, while older universities have to abandon old labs to create new spaces.

**What are the challenges?**

Even in a relatively young country such as the United States, most of its well-known highly ranked universities have been around for a hundred years or more. There have been many experiments around the world to establish new universities, but most have not been able to make it into the global top 50 or 100, even those with lots of funding.

Many universities whose glory days are in the past are still highly ranked. It takes a long time for word to get around that you have reached your peak. This time lag applies in both directions, as it's very difficult for young universities to crack the rankings. But NTU has consistently delivered, and now word is getting around. This year more than 430 papers were published by NTU faculty in the top ten journals in the world.

It's difficult for young universities to compete with well-established institutions, but they see the value of partnering with us. We have strong partnerships with Massachusetts Institute of Technology (MIT) and Imperial College London, and we have a very strong partnership with the Technical University of Munich, Germany, in the area of robotics. In 2018, the Wallenberg Foundation of Sweden gave an endowment to NTU to support postdoctoral researchers.

**How do global metrics affect your strategies?**

We look at all of them, and at other metrics, such as where our faculty publish, the quality of the faculty we recruit and where they come from. In the past 18 months, for example, we've recruited from the University of Cambridge in the UK, American Ivy League schools and from top institutions in Europe and Asia.

In January 2018, a couple of months after I started as president, we launched the presidential postdoctoral fellows programme to attract the brightest young postdocs. This year we had 894 applications from 74 countries for 12 positions.

About a year ago, the US Institute of Electrical and Electronics Engineers listed the top ten rising stars in artificial intelligence around the world. According to their assessment, three are NTU faculty.

Good performance in a ranking can be a motivator, but one cannot take it as the only metric and the only reason to do well.

**What are your priorities from here?**

We are doing our best to attract top talent from Singapore and from all over the world. This includes students, postdocs, faculty and staff. Our commitment to excellence in education and research comes second. It's not research versus education. The two have to be integrated. To have impact we need to make sure that research and education connect with both societal and industrial impact. That's why government and industry partnerships are so important.

Some demographic trends will affect all of the universities in Singapore in the next ten to 15 years. Our birth rate has been declining for many years, and because funding is tied to undergraduate student involvement, this decline will affect us all. We have an obligation to deliver value for the resources we get from the Singapore government. As a young university, we had access to significant new resources to help grow the university, but this upwards trajectory cannot be sustained forever.

We have significant momentum and we will continue to grow in stature and impact in our output in education, research and innovation, but that doesn't mean that the level of annual increase in funding will be the same over the next ten years. So, one priority we have is to continue our growth in excellence without necessarily continuing to grow in numbers.

**Interview by Catherine Armitage**

This interview has been edited for clarity and length.

HUTTON+CROW/VIEW PICTURES/UNIVERSAL IMAGES GROUP VIA GETTY IMAGES